

09/673274

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: October 16, 2003, 17:06:33 ; Search time 8.77076 Seconds  
(without alignments)  
1415.504 Million cell updates/sec

Title: US-09-673-274B-2\_COPY\_1\_132  
Perfect score: 267  
Sequence: 1 gataagcttaccgttctctg.....tgaactcgtggtcgagact 132

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 235726  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool/US09673274/runat\_16102003\_170627\_2531/app\_query.fasta.1.654  
-DB=SwissProt\_41 -FORMAT=fastan -SUFFIX=resp -MINMATCH=0.1 -LOOPEXT=0 -LIST=45  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=expct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09673274@cgn\_1\_1\_36@runat\_16102003\_170627\_2531 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEVO TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6  
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt\_41:.\*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	267	100.0	44	1 DEFN_HELVI	P81544 heliothis v
2	86	32.2	66	1 SCX3_BUTOM	P13488 buthus occi
3	84.5	31.6	85	1 SCN2_MESMA	Q9BKJ1 mesobuthus
4	84	31.5	85	1 SCAS_MESMA	Q9UAC9 mesobuthus
5	83.5	31.3	85	1 SCN3_MESMA	Q9BKJ0 mesobuthus
6	79.5	29.8	66	1 SCXC_CENLL	P45667 centruoide
7	79	29.6	85	1 SCXE_BUTOC	Q17254 buthus occi
8	78.5	29.4	65	1 SCX4_BUTOM	P59354 buthus occi
9	78.5	29.4	84	1 SCX3_ANDAU	P01480 androctonus
10	77.5	29.0	64	1 SCLE_LEIOH	P59356 leiorus qui
11	77.5	29.0	65	1 SCX4_LEIOU	P01489 leiorus qui
12	77	28.8	85	1 SCAT_MESMA	Q9UAC8 mesobuthus
13	76.5	28.7	61	1 SIX2_LEIOU	P19855 leiorus qui
14	76	28.5	66	1 SCXE_BUTEU	P09982 buthus eupe
15	76	28.5	67	1 SCL3_LEIOH	P56678 leiorus qui
16	75.5	28.3	84	1 SCXA_MESMA	O61705 mesobuthus
17	74.5	27.9	70	1 DMYC_DROME	P41964 drosophila
18	74.5	27.9	87	1 SCX1_CENSC	P01492 centruoide

19	74	27.7	59	1 SCX5_CENSC	P58779 centruoide
20	74	27.7	84	1 SCX4_ANDAU	P45658 androctonus
21	73	27.3	61	1 SIX5_ANDAU	P81504 androctonus
22	73	27.3	64	1 SCX5_LEIOU	P01481 leiorus qui
23	72	27.0	85	1 SC15_MESMA	Q9GNG8 mesobuthus
24	72	27.0	85	1 SCX1_MESMA	Q9GXX2 mesobuthus
25	71.5	26.8	61	1 SIX2_BUTAR	P80362 buthacus ar
26	71.5	26.8	61	1 SIX4_BUTOC	P55903 buthus occi
27	71.5	26.8	61	1 SIX5_BUTOC	P55904 buthus occi
28	71.5	26.8	87	1 SCX2_CENSC	P01493 centruoide
29	71.5	26.8	87	1 SCX3_CENSC	P01494 centruoide
30	71	26.6	43	1 DEF1_PALPR	P80407 palomena pr
31	71	26.6	66	1 SCL7_LEIOH	P59357 leiorus qui
32	70.5	26.4	64	1 SCM2_MESMA	P58488 mesobuthus
33	70.5	26.4	84	1 SC16_MESMA	Q9GQV6 mesobuthus
34	70.5	26.4	87	1 SCX6_CENNO	P45664 centruoide
35	69.5	26.0	84	1 SCX1_MESMA	P45697 mesobuthus
36	69.5	26.0	84	1 SCX2_CENNO	P01495 centruoide
37	69.5	26.0	84	1 SCX9_CENSC	Q95WC9 centruoide
38	69.5	26.0	85	1 SC10_MESMA	Q9N1C5 mesobuthus
39	68.5	25.7	62	1 SIX4_BUTSI	P82814 buthus sind
40	68.5	25.7	65	1 SCX1_BUTOC	P01488 buthus occi
41	68.5	25.7	66	1 SCX6_ANDAU	P56743 androctonus
42	68.5	25.7	66	1 SCX9_BUTEU	P09981 buthus eupe
43	68.5	25.7	74	1 DEF2_CAPAN	O65740 capsicum an
44	68.5	25.7	74	1 SC17_MESMA	Q9N1C4 mesobuthus
45	68.5	25.7	83	1 SCX1_ANDAU	P01479 androctonus

ALIGNMENTS

RESULT: 1  
ID DEFN\_HELVI STANDARD; PRT; 44 AA.  
AC P81544;  
DC 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DI 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Defensin heliothis.  
OS Heliothis virescens (Noctuid moth) (Owlet moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;  
OC Noctuidae; Heliothinae; Heliothis.  
OX NCBI\_TaxId=7102;  
RN [1]  
RP SEQUENCE, AND CHARACTERIZATION.  
RC TISSUE=Hemolymph;  
RX MEDLINE=99194775; PubMed=10092609;  
RA Lamberty M., Ades S., Uttenweiler-Joseph S., Brookhart G., Bushey D.,  
RA Hoffmann J.A., Bulet P.;  
RT "Insect immunity. Isolation from the lepidopteran Heliothis virescens  
of a novel insect defensin with potent antifungal activity.";  
RL J. Biol. Chem. 274:9320-9326(1999).  
CC -!- FUNCTION: THIS PEPTIDE HAS POTENT ANTI-FUNGAL ACTIVITY. HAS NC  
CC -!- SUBCELLULAR LOCATION: Secreted  
CC -!- MASS SPECTROMETRY: MW=4784.8; METHOD=MAINT.  
DR PDB; 1I2U; 12-FEB-02.  
DR PDB; 1I2V; 12-FEB-02.  
KW Insect immunity; Fungicide; 3D-structure.  
FT DISULFID 7 32  
FT DISULFID 18 40  
FT DISULFID 22 42  
SQ SEQUENCE 44 AA; 4790 MW; 30-DA1BD3376FCD5 CRC64;

Alignment Scores:  
Pred. No.: 3.22e-27  
Score: 267.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 1  
Length: 44  
Matches: 44  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0



28-FEB-2003 (Rel. 41, Last annotation update)  
 Neurotoxin AS precursor.  
 Mesobuthus martensii (Manchurian scorpion) (Butus martensii).  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 Buthoidea; Buthidae; Mesobuthus.  
 NCBI\_TaxID=34649;  
 (1)  
 SEQUENCE FROM N.A.  
 TISSUE=Venom gland;  
 Zhang J.-H., Hua Z.C., Zhu D.X.;  
 "Cloning of anti-epilepsy peptide cDNA from scorpion Buthus martensii  
 Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases."  
 (2)  
 SEQUENCE FROM N.A.  
 TISSUE=Venom gland;  
 Zhang J.-H., Hua Z.C., Zhu D.X.;  
 "Cloning of anti-neuroexcitation peptide III (ANEP) cDNA from scorpion  
 Buthus martensii Karsch."  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 -!- FUNCTION: Binds to sodium channels and inhibits them.  
 -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
 -!- TISSUE SPECIFICITY: Expressed by the venom gland.  
 -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
 -!- ALPHA-TOXIN SUBFAMILY.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 EMBL; AF079060; AAD47374.1; -;  
 HSSP; P01492; 1VNB.  
 InterPro; IPR003614; Kntol.  
 InterPro; IPR002061; Scorpion\_toxinL.  
 Pfam; PF00537; toxin\_3; 1.  
 ProDom; PD000908; Scorpion\_toxinL; 1.  
 SMART; SM00505; Kntol; 1.  
 Toxin; Neurotoxin; Signal.  
 SIGNAL  
 CHAIN 1 19 POTENTIAL.  
 DISULFID 20 85 NEUROTOXIN AS.  
 BY SIMILARITY.  
 DISULFID 31 81 BY SIMILARITY.  
 DISULFID 35 56 BY SIMILARITY.  
 DISULFID 42 63 BY SIMILARITY.  
 DISULFID 46 65 BY SIMILARITY.  
 DISULFID 47 65 BY SIMILARITY.  
 SEQUENCE 85 AA; 9759 MW; 41286F9DCD4FE90 CRC64;  
 Alignment Scores:  
 Pred. No.: 0.00392 Length: 85  
 Score: 84.00 Matches: 19  
 Percent Similarity: 55.81% Conservative: 5  
 Best Local Similarity: 44.19% Mismatches: 17  
 Query Match: 31.46% Indels: 2  
 DB: 1 Gaps: 2  
 US-09-673-274b-2\_copy\_1\_132 (1-132) x SCAS\_MESMA (1-85)  
 QY 1 GATAAGCTTATCGGTCTCGTGGGGTGTGTAACACTTCCGATTGCAACGGT 60  
 DDb 26 AplysTyrThrGlyCysValTrpCysValIleAsnAsnGluSer---CysAsnSer 44  
 QY 61 GAGTCACAGAGAGGGGTACACAGGGTGGTCACTCGGTTCTTCGCTAAAGTGAACCTGC 120  
 DDb 45 GluCysValIleArgGlyGlyTyrTyrGlyTyrCys---TyrPheTyrLysLeuAlaCys 63  
 QY 121 TGGTCGCGAG 129  
 DDb 64 PheCysGln 66  
 SCNS3\_MESMA  
 ID SCNS3\_MESMA STANDARD; PRT; 85 AA.  
 AC Q9BKJ0; Q9GZC4;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Anti-neuroexcitation peptide III precursor (ANEP111) (Anti-epilepsy

peptide).  
 Mesobuthus martensii (Manchurian scorpion) (Butus martensii).  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 Buthoidea; Buthidae; Mesobuthus.  
 NCBI\_TaxID=34649;  
 (1)  
 SEQUENCE FROM N.A.  
 TISSUE=Venom gland;  
 Zhang J.-H., Hua Z.C., Zhu D.X.;  
 "Cloning of anti-epilepsy peptide cDNA from scorpion Buthus martensii  
 Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases."  
 (2)  
 SEQUENCE FROM N.A.  
 TISSUE=Venom gland;  
 Zhang J.-H., Hua Z.C., Zhu D.X.;  
 "Cloning of anti-neuroexcitation peptide III (ANEP) cDNA from scorpion  
 Buthus martensii Karsch."  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 -!- FUNCTION: Binds to sodium channels and inhibits them.  
 -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
 -!- TISSUE SPECIFICITY: Expressed by the venom gland.  
 -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
 -!- ALPHA-TOXIN SUBFAMILY.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 EMBL; AF122003; AAG01572.1; -;  
 HSSP; P01494; 2SN3.  
 InterPro; IPR003614; Kntol.  
 InterPro; IPR002061; Scorpion\_toxinL.  
 Pfam; PF00537; toxin\_3; 1.  
 ProDom; PD000909; Scorpion\_toxinL; 1.  
 SMART; SM00505; Kntol; 1.  
 Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;  
 SIGNAL  
 CHAIN 1 21 POTENTIAL.  
 DISULFID 22 85 ANTI-NEUROEXCITATION PEPTIDE III.  
 BY SIMILARITY.  
 DISULFID 31 81 BY SIMILARITY.  
 DISULFID 35 56 BY SIMILARITY.  
 DISULFID 42 63 BY SIMILARITY.  
 DISULFID 46 65 BY SIMILARITY.  
 DISULFID 47 65 BY SIMILARITY.  
 SEQUENCE 85 AA; 927 MW; C1D6C91A2F82F8C2 CRC64;  
 Alignment Scores:  
 Pred. No.: 0.00457 Length: 85  
 Score: 83.50 Matches: 16  
 Percent Similarity: 55.26% Conservative: 5  
 Best Local Similarity: 42.11% Mismatches: 12  
 Query Match: 31.27% Indels: 5  
 DB: 1 Gaps: 2  
 US-09-673-274b-2\_COPY\_1\_132 (1-132) x SCN3\_MESMA (1-85)  
 QY 16 TCCTGGGTGGGGTGTGTAACACTTCCGATTGCAACGGTGAAGAGAGAGG 75  
 DDb 34 SerCysSerTrpGly-----AsnGluGlyCysAsnLysGluCysLysGlyPhe 49  
 QY 76 GGTACAGAGGGTGTCACTCGGTTCTTCGCTAAAGTGAACCTGCCTGTCGGAG 129  
 DDb 50 GlyAlaTyrTyrGlyTyrCysTrpTrpTrp---GlyLeuAlaCysTrpCysGlu 64  
 RESULT 6  
 SCXC\_CENLL  
 ID SCXC\_CENLL STANDARD; PRT; 66 AA.

AC P45667;  
 DT 01-NOV-1995 (Rel. 32, Created);  
 DT 01-NOV-1995 (Rel. 32, Last sequence update);  
 DT 28-FEB-2003 (Rel. 41, Last annotation update);  
 DE Clustacean-specific toxin 1;  
 DE Centruroides limpidus limpidus (Mexican scorpion);  
 CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 CC Buthidae; Buthidae; Centruroides;  
 CC NCBI\_TaxID=29941;  
 RN 1;  
 RP SEQUENCE, DISULFIDE BONDS, AND STRUCTURE BY NXR;  
 RC 155JRH:venom;  
 RC MEDLINE=95244424; PubMed=7127365;  
 RA Lebreton F., Delapierre M., Ramirez A.N., Balderas C., Possani L.D.;  
 RA Primarity and NMR three-dimensional structure determination of a novel  
 PT clustacean toxin from the venom of the scorpion Centruroides limpidus  
 RT limpidus Karsch;  
 RT Biochemistry 33:11135-11149 (1994);  
 CC FUNCTION: Binds to sodium channels and inhibits the inactivation  
 CC of the activated channels, thereby blocking neuronal transmission.  
 CC SUBCELLULAR LOCATION: Secreted.  
 CC TISSUE SPECIFICITY: Expressed by the venom gland.  
 CC SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
 CC BETA-TOXIN SUBFAMILY.  
 CC PIR: A55869; A55869;  
 CC HSSP: F01492; 1VNB;  
 CC InterPro: IPR003614; Knott1;  
 CC InterPro: IPR001219; Neurotoxin;  
 CC InterPro: IPR002061; Scorpion\_toxinL;  
 CC Pfam: PF00537; toxin\_3; 1;  
 CC ProDom: PD003908; Scorpion\_toxinL; 1;  
 CC PRINTS: PR0284; TOXIN;  
 CC PROSITE: PS00908; Scorpion\_toxinL; 1;  
 CC SMART: SM00505; Knott1; 1;  
 CC TIGR: Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;  
 FT DISULFID 12 65  
 FT DISULFID 16 41  
 FT DISULFID 25 46  
 FT DISULFID 29 48  
 SQ SEQUENCE 66 AA; 7338 MW; A478E63796F54DAA CNC64;  
 Alignment Scores:  
 Pred. No.: 0 C:51 Length: 66  
 Score: 79.50 Matches: 17  
 Percent Similarity: 53.85% Conservative: 4  
 Best Local Similarity: 43.59% Mismatches: 11  
 Query Match: 29.78% Indels: 7  
 Gaps: 3  
 DB: 1  
 US-09-673-274b-2\_COPY\_1\_132 (1-132) x SCXC\_CENLL (1-66)  
 QY 19 TCGTGTGGGGTCTGTGAACACTACACTTCGATTGCAACGGTGAGTGCAAGAGGGGGT 78  
 Db 16 CysPheTrpLeuGlyLysAsn-----GluAsnCysAspLysLysLysAla-LysAsn 33  
 QY 79 TACAAGGGT-----CGTCACTCGGTTCTCTCGTTAACTGAACGTGGTGGCGAG 129  
 DE 34 GlnGlyGlySerTyrG.YTyrCysTyrSerPheAla-----CysTrpCysGlu 49  
 RESULT 7  
 SCXE\_BUTOC STANDARD; PRT; 85 AA.  
 ID SCXE\_BUTOC  
 AC Q17254;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Neurotoxin XIV precursor (Bot14);  
 DE Suchus occitanus tunicatus (Common European scorpion);  
 CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 CC Buthidae; Buthidae;  
 CC NCBI\_TaxID=6871;  
 RN 1;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE:Venom;

RX MEDLINE=96300228; PubMed=8706664;  
 RA Bouchouala-Zahar B., Ducancel F., Zenouaki J., Ben Khalifa R.,  
 RA Borchani L., Pelhate M., Boulain J.C., el Ayeub M., Menez A.,  
 RA Karoui H.;  
 RT "A recombinant insect-specific alpha-toxin of Buthus occitanus  
 RT tunicatus scorpion confers protection against homologous mammal  
 RT toxins";  
 RL Eur. J. Biochem. 238:653-660 (1996).  
 CC FUNCTION: Binds to sodium channels and inhibits the inactivation  
 CC of the activated channels, thereby blocking neuronal transmission.  
 CC This toxin is active only on insects.  
 CC SUBCELLULAR LOCATION: Secreted.  
 CC TISSUE SPECIFICITY: Expressed by the venom gland.  
 CC SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
 CC ALPHA-TOXIN SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X92376; CAA63120.1;  
 CC PIR: S68906; S68906;  
 CC HSSP: P56678; 1BMR;  
 CC InterPro: IPR003614; Knott1;  
 CC InterPro: IPR002061; Scorpion\_toxinL;  
 CC Pfam: PF00537; toxin\_3; 1;  
 CC ProDom: PD003908; Scorpion\_toxinL; 1;  
 CC SMART: SM00505; Knott1; 1;  
 CC TIGR: Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;  
 KW SIGNAL;  
 FT CHAIN 19 85 POTENTIAL;  
 FT DISULFID 30 83 BY SIMILARITY;  
 FT DISULFID 34 55 BY SIMILARITY;  
 FT DISULFID 41 65 BY SIMILARITY;  
 FT DISULFID 45 67 BY SIMILARITY;  
 SQ SEQUENCE 85 AA; 9174 MW; 4977F16C511F0567 CRC64;  
 Alignment Scores:  
 Pred. No.: C:0178 Length: 85  
 Score: 79.00 Matches: 14  
 Percent Similarity: 50.00% Conservative: 6  
 Best Local Similarity: 35.00% Mismatches: 18  
 Query Match: 29.59% Indels: 2  
 Gaps: 1  
 DB: 1  
 US-09-673-274b-2\_COPY\_1\_132 (1-132) x SCXE\_BUTOC (1-85)  
 QY 16 TCCTGCGTGGGGTGTGTGAACACTACACTTCGATTGCAACGGTGAGTGCAAGAGGAGG 75  
 Db 29 AsnCysAlaTyrHisCysLeuLysLysLysLysSerSerGlyCysAepThrLeuCysLysGluAsn 48  
 QY 76 GGTTCACAGGGTGTCTACTCGGT-----TCCTTCGCTAACGTGAACGTCTGCTGGAG 129  
 Db 49 GlyAlaThrSerGlyHisCysG.YHisLysSerGlyHisGlySerAlaCysTrpCysLys 68  
 RESULT 8  
 SCX4\_BUTOM STANDARD; PRT; 65 AA.  
 ID SCX4\_BUTOM  
 AC P59154;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Alpha-like neurotoxin IV (Bom IV) (Bom4)  
 DE Suchus occitanus mardochei (Moroccan scorpion);  
 CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 CC Buthidae; Buthidae;  
 CC NCBI\_TaxID=6869;  
 RN 1;  
 RP TISSUE:Venom;



SEQUENCE, PHARMACOLOGICAL CHARACTERIZATION, AND MASS SPECTROMETRY.  
TISSUE=Venom;  
MEDLINE=99203324; PubMed=10103091;  
Cestele S., Stankiewicz M., Mansuelle P., De Waard M., Dargent B.,  
Gilles N., Pelhate M., Rochat H., Martin-Eaucalire M.-F., Gordon D.,  
"Scorpion alpha-like toxins, toxic to both mammals and insects,  
differentially interact with receptor site 3 on voltage-gated sodium  
channels in mammals and insects."  
Eur. J. Neurosci. 11:975-985(1999).  
-1- FUNCTION: Binds to sodium channels and inhibits the inactivation  
of the activated channels, thereby blocking neuronal transmission.  
-1- SUBCELLULAR LOCATION: Secreted.  
-1- TISSUE SPECIFICITY: Expressed by the venom gland.  
-1- MASS SPECTROMETRY: MW=7287.96; MW ERR=0.37; METHOD=Electrospray.  
-1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
-1- ALPHA-TOXIN SUBFAMILY: ALPHA-LIKE TOXIN SUBCLASS.  
Pfam: PF00537; toxin\_3; 1.  
ProDom: PD000908; Scorpion\_toxinlb; 1.  
SMART: SM00505; Kntcl; 1.  
Toxin: Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.  
DISULFID 12 63 BY SIMILARITY.  
DISULFID 16 36 BY SIMILARITY.  
DISULFID 22 46 BY SIMILARITY.  
DISULFID 26 48 BY SIMILARITY.  
SEQUENCE 65 AA; 7296 MW; 4305EB32436A4B90 CRC64;  
Alignment Scores:  
Pred. No.: 0.0204 Length: 65  
Score: 78.50 Matches: 16  
Percent Similarity: 51.22% Conservative: 5  
Best Local Similarity: 39.02% Mismatches: 15  
Query Match: 29.40% Indels: 5  
DB: 1 Gaps: 3  
US-09-673-274b-2\_COPY\_1\_132 (1-132) x SCX4\_BUTOM (1-65)  
16 TCCTCGGTGGGGTGGTGTGCAACTTCCGATTGCAACGGTGAGTGAAGAGGAGG 75  
11 AsnCysValTyrGluCysAlaLys----AsnSerTyrCysAsnAspLeuCyThrLysAsn 29  
76 GGTACAAAGGGTGTCACTGC-----GGTTCCTTCGTAACGTGAACGTGCTGTGC 126  
30 GlyAlaLysSerGlyTyrCysGlnTrpLeuGlyLysTyrGlyAsnAla---CysTrpCys 48  
127 GAG 129  
49 Glu 49  
ID SCX3 ANDAU STANDARD; PRT; 84 AA.  
SEQUENCE FROM N.A.  
MEDLINE=90037062; PubMed=2808423;  
Bougis P.E., Rochat H., Smith L.A.;  
"Precursors of Androctonus australis scorpion neurotoxins. Structures  
of precursors, processing outcomes, and expression of a functional  
recombinant toxin II."  
J. Biol. Chem. 264:19259-19265(1989).  
SEQUENCE OF 20-83.  
MEDLINE=79148414; PubMed=428402;  
Kopeyan C., Martinez G., Rochat H.;  
"Amino acid sequence of neurotoxin III of the scorpion Androctonus

australis Hector."  
Eur. J. Biochem. 94:609-615(1979).  
STRUCTURE BY NMR.  
MEDLINE=91122921; PubMed=2279844;  
Laplante S.R., Mikou A., Robin M., Guittet E., Delsuc M.A.,  
Charpentier J., Lallemand J.-Y.;  
"Rapid determination and NMR assignments of antiparallel sheets and  
helices of a scorpion and a cobra toxin."  
Int. J. Pept. Protein Res. 36:227-230(1990).  
STRUCTURE BY NMR.  
MEDLINE=93044377; PubMed=1422146;  
Mikou A., Laplante S.R., Guittet E., Lallemand J.-Y.,  
Martin-Eaucalire M.-F., Rochat H.;  
"Toxin III of the scorpion Androctonus australis Hector: proton  
nuclear magnetic resonance assignments and secondary structure."  
J. Biomol. NMR 2:57-70(1992).  
-1- FUNCTION: Binds to sodium channels and inhibits the inactivation  
of the activated channels, thereby blocking neuronal transmission.  
-1- SUBCELLULAR LOCATION: Secreted.  
-1- TISSUE SPECIFICITY: Expressed by the venom gland.  
-1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
-1- ALPHA-TOXIN SUBFAMILY.  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation.  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL; M27703; AAA29948.1;  
P.R; C34444; NTSR3A.  
HSP; P56678; 15MR.  
InterPro: IPR03614; Kntcl.  
Pfam: PF00537; toxin\_3; 1.  
SMART: PD000908; Scorpion\_toxinlb; 1.  
Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.  
Signal.  
CHAIN 1 19 NEUROTOXIN III  
PROPEP 20 83  
DISULFID 31 81 BY SIMILARITY.  
DISULFID 35 53 BY SIMILARITY.  
DISULFID 39 63 BY SIMILARITY.  
DISULFID 43 65 BY SIMILARITY.  
CONFLICT 27 27 D -> N (IN REF. 2).  
SEQUENCE 84 AA; 9068 MW; ADB0A3462153E3C0 CRC64;  
Alignment Scores:  
Pred. No.: 0.0207 Length: 84  
Score: 78.50 Matches: 15  
Percent Similarity: 53.85% Conservative: 6  
Best Local Similarity: 38.46% Mismatches: 13  
Query Match: 29.40% Indels: 5  
DB: 1 Gaps: 2  
US-09-673-274b-2\_COPY\_1\_132 (1-132) x SCX3\_ANDAU (1-84)  
Qy 16 TCCTCGGTGGGGTGGTGTGCAACTTCCGATTGCAACGGTGAGTGAAGAGGAGG 75  
Db 30 AsnCysValTyrHisCysVal-----ProProCysAspGlyLeuCysLysAsn 46  
Qy 76 GGTACAAAGGGTGTCACTGCCTCCGTTCTCTTC-----GCTAACGTGAACGTGCTGTGC 126  
Db 47 GlyAlaLysSerGlySerCysGlyPheLeuLeuProSerGlyLeuAlaCysTrpCys 65  
RESULT 10  
SCL6\_LEIQH

SCX4\_LEIQH STANDARD; PRT; 64 AA.  
 AC P59156;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Alpha-like neurotoxin Lqn VI (Lqn6).  
 CS Leirus quinquestriatus hebraeus (yellow scorpion).  
 CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 CC Butthidea; Butthidae; Leiurus.  
 CC NCBI\_TaxID=6884;  
 RN [1]  
 RP SEQUENCE, PHARMACOLOGICAL CHARACTERIZATION, AND BIOASSAY.  
 RC TISSUE=Venom;  
 RA MEDLINE=22169893; PubMed=12180969;  
 RA Hamon A., Gilles N., Sautiere P., Martinage A., Kopeyan C., Ulens C.,  
 RA Tytgat C., Lancelin J.-M., Gordon D.;  
 RT Characterization of scorpion alpha-like toxin group using two new  
 RT toxins from the scorpion Leiurus quinquestriatus hebraeus.;  
 RJ Eur. J. Biochem. 269:3920-3933(2002).  
 CC !- FUNCTION: Binds to sodium channels and inhibits the inactivation  
 CC of the activated channels, thereby blocking neuronal transmission.  
 CC This toxin is highly toxic to insects and mice, and inhibits the  
 CC binding of alpha-toxin to cockroach neuronal membranes.  
 CC !- SUBCELLULAR LOCATION: Secreted.  
 CC !- TISSUE SPECIFICITY: Expressed by the venom gland.  
 CC ! MISCELLANEOUS: LD50 is 34.3 nmol/kg to cockroaches (Blattella  
 CC germanica).  
 CC !- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
 CC ALPHA-TOXIN SUBFAMILY. ALPHA-LIKE TOXIN SUBCLASS.  
 DR Pfam: PF00537; toxin\_3; 1.  
 DR ProDom: PD000908; Scorpion\_toxinL; 1.  
 DR SMART: SM00505; Krot1; 1.  
 DR Toxin: Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;  
 KW Amidation.  
 FT DISULFID 12 62 BY SIMILARITY.  
 FT DISULFID 16 34 BY SIMILARITY.  
 FT DISULFID 20 44 BY SIMILARITY.  
 FT DISULFID 24 46 BY SIMILARITY.  
 FT MOD\_RES 64 64 AMIDATION.  
 FT MOD\_RES 64 64 AMIDATION.  
 SQ SEQUENCE 64 AA; 6803 MW; 49A54BCB6C75392E CRC64;  
 Alignment Scores:  
 Pred. No.: 0.0276 Length: 64  
 Score: 77.50 Matches: 15  
 Percent Similarity: 51.28% Conservative: 5  
 Best Local Similarity: 38.46% Mismatches: 14  
 Query Match: 29.03% Indels: 5  
 DB: 1 Gaps: 2  
 US-09-673-274b-2\_copy\_1\_132 (1-132) x SCL6\_LEIQH (1-64)  
 QY 16 TCCTGGTGTGGGTGCTGTGAACACTTCCGATTGCAACGGTGAAGTCAAGAGGAGG 75  
 DB 11 AsnCysValtyrHisCysIle-----ProAspCysAspThrLeuCysLysAspAsn 27  
 QY 76 GGTACAAAGGGTGGTCACTCGGTTCTTCGCT-----AACGTGAACCTGGTGGTCC 126  
 DB 28 GlyGlyThrGlyGlyHisCysGlyPheLeuLysLeuGlyHisGlyIleAlaCysTrpCys 46  
 RESULT 11  
 SCX4\_LEIQH  
 ID SCX4\_LEIQH STANDARD; PRT; 65 AA.  
 AC P01489;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Neurotoxin IV (Lqn IV) (Lqn4).  
 CS Leiurus quinquestriatus quinquestriatus (Egyptian scorpion).  
 CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 CC Butthidea; Butthidae; Leiurus.  
 CC NCBI\_TaxID=6885;  
 RN [1]  
 RP SEQUENCE.

RA Kopeyan C., Martinez G., Rochat H.;  
 RT "Primary structure of toxin IV of Leiurus quinquestriatus  
 RL quinquestriatus.";  
 RL FEBS Lett. 181:211-217(1985).  
 CC !- FUNCTION: Binds to sodium channels and inhibits the inactivation  
 CC of the activated channels, thereby blocking neuronal transmission.  
 CC !- SUBCELLULAR LOCATION: Secreted.  
 CC !- TISSUE SPECIFICITY: Expressed by the venom gland.  
 CC !- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
 CC ALPHA-TOXIN SUBFAMILY.  
 DR PIR: A01749; NTSR4L.  
 DR HSP: P45697; ISN1.  
 DR InterPro: IPR003614; Krot1.  
 DR InterPro: IPR001219; Neurotoxin.  
 DR InterPro: IPR002061; Scorpion\_toxinL.  
 DR Pfam: PF00537; toxin\_3; 1.  
 DR PRINTS: PR00284; TCXIN.  
 DR ProDom: PD000908; Scorpion\_toxinL; 1.  
 DR SMART: SM00505; Krot1; 1.  
 DR Toxin: Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;  
 KW Amidation.  
 FT DISULFID 13 64 BY SIMILARITY.  
 FT DISULFID 17 37 BY SIMILARITY.  
 FT DISULFID 23 47 BY SIMILARITY.  
 FT DISULFID 27 49 BY SIMILARITY.  
 FT MOD\_RES 65 65 AMIDATION.  
 FT MOD\_RES 65 65 AMIDATION.  
 SQ SEQUENCE 65 AA; 7202 MW; 882102EC1C2B89F6 CRC64;  
 Alignment Scores:  
 Pred. No.: 0.0276 Length: 65  
 Score: 77.50 Matches: 16  
 Percent Similarity: 57.14% Conservative: 8  
 Best Local Similarity: 38.10% Mismatches: 9  
 Query Match: 29.03% Indels: 9  
 DB: 1 Gaps: 4  
 US-09-673-274b-2\_copy\_1\_132 (1-132) x SCX4\_LEIQH (1-65)  
 QY 16 TCCTGGTGTGGTGG-----GGTGTGTGAACACTTCCGATTGCAACGGTGAAGTCAAG 69  
 DB 12 AsnCysValtyrThrCysGlySerAsnSerTyr-----CysAsnThrGlyCysThr 28  
 QY 70 AGAGGGGTTCACAAAGGGTGGTCACTGC-----GGTTCCTTCGCTAACGTGAAGTGC 120  
 DB 29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpLeuGlyLysThrGlyAsnAla---Cys 47  
 QY 121 TGGTGC 126  
 DB 48 TrpCys 49  
 RESULT 12  
 SCAT\_MESMA  
 ID SCAT\_MESMA STANDARD; PRT; 85 AA.  
 AC Q9UAC8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Neurotoxin AS-1 precursor.  
 OS Mesobuthus martensii (Menchurian scorpion) (Buthus martensii).  
 CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 CC Butthidea; Butthidae; Mesobuthus.  
 CC NCBI\_TaxID=34649;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Venom gland;  
 RA Lan Z.D., Dai L., Chi C.-W.;  
 RT "The gene cloning and sequencing of AS and AS-1, two novel neurotoxins  
 RT from the scorpion Buthus martensii Karsch.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 CC !- FUNCTION: Ryanodine receptor agonist in skeletal muscle.  
 CC !- SUBCELLULAR LOCATION: Secreted (by similarity).  
 CC !- TISSUE SPECIFICITY: Expressed by the venom gland.  
 CC !- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.

ALPHA-TOXIN SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

EMBL; AF079061; AAD7475.1; -

HSSP; P01492; 1VNB.

InterPro; IPR003614; Scorpion\_toxinL.

InterPro; IPR002061; Scorpion\_toxinL.

Pfam; PF00537; toxin\_3; 1.

ProDom; PD000908; Scorpion\_toxinL; 1.

SMART; SM00505; Knot1; 1.

Toxin; Neurotoxin; Signal.

SIGNAL 1 19

POTENTIAL.

CHAIN 20 85

NEUROTOXIN AS-1.

DISULFID 31 81

BY SIMILARITY.

DISULFID 35 56

BY SIMILARITY.

DISULFID 42 63

BY SIMILARITY.

DISULFID 46 65

BY SIMILARITY.

SEQUENCE 85 AA; 9802 MW; E2274838439E2B95 CRC64;

Alignment Scores:

Pred. No.:	0.0327	Length:	85
Score:	77.00	Matches:	17
Percent Similarity:	53.49%	Conservative:	6
Best Local Similarity:	39.53%	Mismatches:	18
Query Match:	28.84%	Indels:	2
DB:	1	Gaps:	2

US-09-673-274b-2\_COPY\_1\_132 (1-132) x SCAT\_MESMA (1-85);

QY 1 GATAGCTTATCGGTCTCGGTGGGTGCTGGAACACTACCTTCGGATTGCAACGGT 60

Db 26 AsnLysTyrThrGlyCysAlaLeuTrpCysValleAsnAsnGluSer---CysAsnSer 44

QY 61 GAGTCAAGAGAGGGGTACCAAGGGTGGTCACTGGCGTCTCTCGCTAAGCTGAACCTGC 120

Db 45 GluCysLysLeuArgGlyAsnTyrGlyTyrCys---TyrPheTrpLysLeuAlaCys 63

QY 121 TCGTCCGAG 129

Db 64 TyrCysGlu 66

RESULT 13

SIX2 LEIQU STANDARD; PRT; 61 AA.

AC P19855;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Insect toxin 2 (Insect toxin LqgIT2).

OS Leiurus quinquestriatus quinquestriatus (Egyptian scorpion).

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;

OC Buthoidea; Buthidae; Leiurus.

OX NCBI\_TaxID=6885;

RN [1]

RP SEQUENCE.

RC MEDLINE=90184494; PubMed=2111768;

RA Kopyan C., Mansuelle P., Samperi F., Brando T., Bahraoui E.M.,

RA Rochat H., Granier C.

RT "Primary structure of scorpion anti-insect toxins isolated from the venom of Leiurus quinquestriatus quinquestriatus.";

FEBS Lett. 261:423-426(1990).

RN [2]

RP SEQUENCE.

RC MEDLINE=93160530; PubMed=8431601;

RA Zlotkin E., Gurevitz M., Fowler E., Adams M.E.;

RT "Depressant insect selective neurotoxins from scorpion venom:

Chemistry, action, and gene cloning".

Arch. Insect Biochem. Physiol. 22:55-73(1993).

RN [3]

RP ACTION.

RA Zlotkin E., Fowler E., Eitan M., Moyer M., Adams M.E.;

RT "On the chemistry and action of the depressant insect toxins.";

RL Toxicon 28:170-170(1993).

CC -!- FUNCTION: Binds to sodium channels and inhibits the inactivation of the activated channels, thereby blocking neuronal transmission.

CC This toxin induces a slow, depressant, flaccid paralysis on fly larvae. It is active only on insects.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.

CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.

CC ALPHA-TOXIN SUBFAMILY.

DR PIR; B34123; B34123.

DR HSSP; P01494; 2SN3.

DR InterPro; IPR003614; Knot1.

DR InterPro; IPR002061; Scorpion\_toxinL.

DR Pfam; PF00537; toxin\_3; 1.

DR ProDom; PD000908; Scorpion\_toxinL; 1.

DR SMART; SM00505; Knot1; 1.

DR Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.

FT DISULFID 10 60

BY SIMILARITY.

FT DISULFID 14 35

BY SIMILARITY.

FT DISULFID 22 42

BY SIMILARITY.

FT DISULFID 25 44

BY SIMILARITY.

SQ SEQUENCE 61 AA; 6850 MW; 8CF01771C7DCFAF6 CRC64;

Alignment Scores:

Pred. No.:	0.0372	Length:	61
Score:	76.50	Matches:	15
Percent Similarity:	55.26%	Conservative:	6
Best Local Similarity:	39.47%	Mismatches:	12
Query Match:	28.65%	Indels:	5
DB:	1	Gaps:	2

US-09-673-274b-2\_COPY\_1\_132 (1-132) x SIX2\_LEIQU (1-61);

QY 16 TCTCGGTGGGTGCTGGAACACTACCTTCGGATTGCAACGGTGAAGAGAGG 74

Db 13 SerCysLeuPheGly-----AsnGluGlyCysAsnLysGluCysSerTyr 26

QY 76 GGTACAGGGTGGTCACTGGCTGCTTCTCGTAACTGTAACGTGGTGGCGAG 129

Db 29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 14

SCXE BUTEU STANDARD; PRT; 66 AA.

ID SCXE BUTEU

AC P09982;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Neurotoxin M14.

OS Buthus eupeus (Lesser Asian scorpion) (Mesobuthus eupeus).

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;

OC Buthoidea; Buthidae; Mesobuthus.

OX NCBI\_TaxID=34648;

RN [1]

RP SEQUENCE.

RC TISSUE=Venom;

RX MEDLINE=85046632; PubMed=6497916;

RA Volkova T.M., Garsia A.F., Telezhinskaya I.N., Potapenko N.A.,

RA Grishin E.V.;

RT "Amino acid sequence of 2 neurotoxins from the scorpion Buthus eupeus venom.";

RL Bioorg. Khim. 10:979-982(1984).

CC -!- FUNCTION: Binds to sodium channels and inhibits the inactivation of the activated channels, thereby blocking neuronal transmission.

CC Has paralytic activity in mice.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

DM nucleic - protein search, using frame\_plus\_n3p model

Run on: October 16, 2003, 17:06:33 ; Search time 40.7841 Seconds  
(without alignments)  
1027.456 Million cell updates/sec

Title: US-09-673-274B-2\_COPY\_1\_132

Perfect score: 267

Sequence: 1 gataagcttcggtctctg.....tgaaactcgtgcgaact 132

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first: 45 summaries

Command line parameters:

-MODBL=frame\*\_n2p.model -DEV=xlp  
-o=/cgn2\_17/USPTO\_spo01/US09673274/runat\_16102003\_170626\_2523/app\_query.fasta\_1.654  
-DB=A Geneseq\_19Jun03 -QMT=faetan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=0.0 -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEADSIZ=500 -MINLEN=C -MAXLEN=2000000000  
-USER=US09673274@cgn\_1\_1\_195@runat\_16102003\_170626\_2523 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq\_19Jun03:

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	267	100.0	44	21	AAV54375 Amino acid sequenc
2	267	100.0	44	23	ABB76860 Antifungal/antibac
3	267	100.0	44	23	ABB76868 Antifungal/antibac
4	267	100.0	49	21	AAV54374 Amino acid sequenc
5	266	99.6	44	23	ABB76895 Antifungal/antibac
6	264	98.9	44	23	ABB76863 Antifungal/antibac
7	264	98.9	44	23	ABB76864 Antifungal/antibac
8	264	98.9	44	23	ABB76872 Antifungal/antibac
9	264	98.9	44	23	ABB76896 Antifungal/antibac
10	262	98.1	44	23	ABB76862 Antifungal/antibac
11	262	98.1	44	23	ABB76865 Antifungal/antibac
12	262	98.1	44	23	ABB76867 Antifungal/antibac
13	262	98.1	44	23	ABB76869 Antifungal/antibac
14	262	98.1	44	23	ABB76871 Antifungal/antibac
15	262	98.1	74	21	AAV54376 A fusion peptide o
16	261	97.8	44	23	ABB76866 Antifungal/antibac
17	261	97.8	44	23	ABB76870 Antifungal/antibac
18	261	97.8	44	23	ABB76894 Antifungal/antibac
19	259	97.0	44	23	ABB76873 Antifungal/antibac
20	258	96.6	44	23	ABB76898 Antifungal/antibac
21	256	95.9	44	23	ABB76861 Antifungal/antibac
22	256	95.9	44	23	ABB76897 Antifungal/antibac
23	255	95.5	44	23	ABB76882 Antifungal/antibac
24	254	95.1	44	23	ABB76874 Antifungal/antibac
25	254	95.1	44	23	ABB76875 Antifungal/antibac
26	253	94.8	44	23	ABB76879 Antifungal/antibac
27	253	94.8	44	23	ABB76885 Antifungal/antibac
28	251	94.0	44	23	ABB76876 Antifungal/antibac
29	251	94.0	44	23	ABB76877 Antifungal/antibac
30	251	94.0	44	23	ABB76881 Antifungal/antibac
31	251	94.0	44	23	ABB76884 Antifungal/antibac
32	251	94.0	44	23	ABB76886 Antifungal/antibac
33	250	93.6	44	23	ABB76880 Antifungal/antibac
34	250	93.6	44	23	ABB76883 Antifungal/antibac
35	248	92.9	44	23	ABB76887 Antifungal/antibac
36	248	92.9	44	23	ABB76888 Antifungal/antibac
37	246	92.1	44	23	ABB76878 Antifungal/antibac
38	246	92.1	44	23	ABB76889 Antifungal/antibac
39	246	92.1	44	23	ABB76890 Antifungal/antibac
40	243	91.0	44	23	ABB76892 Antifungal/antibac
41	243	91.0	44	23	ABB76899 Antifungal/antibac
42	241	90.3	44	23	ABB76891 Antifungal/antibac
43	238	89.1	44	23	ABB76893 Antifungal/antibac
44	232	86.9	44	24	ABP55988 Agrotis ipsilon fu
45	232	86.9	73	24	ABP55987 Agrotis ipsilon fu

## ALIGNMENTS

RESULT 1  
AAV54375  
In AAV54375 standard; Peptide; 44 AA  
XX  
AC AAV54375;  
XX  
XX C6-APR-2000 (first entry);  
XX  
XX Amino acid sequence of a heliomycin peptide.  
XX  
XX Heliomycin; polycysteine peptide; antifungal; antibacteria.;  
XX fungal infection; plant protection; animal protection;  
XX transgenic plant; fungi; bacteria; Cercospora beticola;  
XX Cladosporium herbarum; Fusarium culmorum; F. oxysporum;  
XX Phytophthora cinnamoni.  
XX  
XX Heliothis virescens.  
OS  
XX





XX 13-JUL-2000; 2000FR-0009248.  
PR 19 SEP 2000; 2000FR-0011949.  
XX (ENTO-) ENTOMED SA.  
PI Dimarcq J, Legrain M, Menin L;  
XX WPI; 2002-179779/23.  
XX The present invention relates to peptide derivatives of heliomicine with  
XX antifungal and/or antibacterial activity. The present sequence is one  
XX such peptide. The peptides are useful in human and veterinary medicine  
XX and in plant protection, for control of bacterial and fungal infections.  
XX An antifungal and/or antibacterial composition containing at least one  
XX peptide is also claimed.  
SQ Sequence 44 AA;  
Alignment Scores:  
Pred. No.: 6,33e-27 Length: 44  
Score: 266.00 Matches: 43  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 97.73% Mismatches: 0  
Query Match: 99.63% Indels: 0  
DB: 23 Gaps: 0  
US-09-673-274b-2\_COPY\_1\_132 (1-132) x ABB76895 (1-44)  
QY 1 GATAAGCTTATCGTTCTCGTGGTGTGCTGCACTACACTCCGATTGCAACGGT 60  
DB : AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20  
QY 61 GAGTGAAGAGGAGGGTTACAAAGGTGGTCACTGCGTTCCTTCGCTAAGTGAACGTC 120  
DB 21 GluCysLysArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnIleAsnCys 40  
QY 121 TGGTGCAGACT 132  
DB 41 TrpCysGluThr 44  
RESULT 6  
ABB76893  
ID ABB76893 standard; Protein: 44 AA.  
XX AC ABB76893;  
XX DT 03-JUL-2002 (first entry)  
XX DE Antifungal/antibacterial peptide PEM38.  
XX KW Antifungal; antibacterial; fungicide; heliomicine; human medicine;  
XX veterinary medicine; plant protection; bacterial infection;  
XX fungal infection.  
XX OS Synthetic.  
XX PN WO200206324-A2.  
XX PD 24 JAN 2002.  
XX PF 05-JUL 2001; 2001WO-FR02164.  
XX PR 13-JUL-2000; 2000FR-0009248.  
XX PR 19-SEP-2000; 2000FR-0011949.  
XX PA (ENTO-) ENTOMED SA.  
XX PI Dimarcq J, Legrain M, Menin L;  
XX WPI; 2002-179779/23.  
XX New peptide derivatives of heliomicine, useful as antibacterial and  
XX antifungal agents in human or veterinary medicine and agriculture -

PI Dimarcq J, Legrain M, Menin L;  
XX WPI; 2002-179779/23.  
XX New peptide derivatives of heliomicine, useful as antibacterial and  
XX antifungal agents in human or veterinary medicine and agriculture -  
XX Claim 15; Page 57; 100pp; French.  
XX The present invention relates to peptide derivatives of heliomicine with  
XX antifungal and/or antibacterial activity. The present sequence is one  
XX such peptide. The peptides are useful in human and veterinary medicine  
XX and in plant protection, for control of bacterial and fungal infections.  
XX An antifungal and/or antibacterial composition containing at least one  
XX peptide is also claimed.  
SQ Sequence 44 AA;  
Alignment Scores:  
Pred. No.: 1.17e-26 Length: 44  
Score: 264.00 Matches: 43  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 97.73% Mismatches: 0  
Query Match: 98.88% Indels: 0  
DB: 23 Gaps: 0  
US-09-673-274b-2\_COPY\_1\_132 (1-132) x ABB76863 (1-44)  
QY 1 GATAAGCTTATCGTTCTCGTGGTGTGCTGAACTACACTCCGATTGCAACGGT 60  
DB 1 AspLysLeuIleGlyThrCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20  
QY 61 GAGTGAAGAGGAGGGTTACAAAGGTGGTCACTGCGTTCCTTCGCTAAGTGAACGTC 120  
DB 21 GluCysLysArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40  
QY 121 TGGTGCAGACT 132  
DB 41 TrpCysGluThr 44  
RESULT 7  
ABB76864  
ID ABB76864 standard; Protein: 44 AA.  
XX AC ABB76864;  
XX DT 03-JUL-2002 (first entry)  
XX DE Antifungal/antibacterial peptide PEM43.  
XX KW Antifungal; antibacterial; fungicide; heliomicine; human medicine;  
XX veterinary medicine; plant protection; bacterial infection;  
XX fungal infection.  
XX OS Synthetic.  
XX PN WO200206324-A2.  
XX PD 24-JAN-2002.  
XX PF 05-JUL-2001; 2001WO-FR02164.  
XX PR 13-JUL-2000; 2000FR-0009248.  
XX PR 19-SEP-2000; 2000FR-0011949.  
XX PA (ENTO-) ENTOMED SA.  
XX PI Dimarcq J, Legrain M, Menin L;  
XX WPI; 2002-179779/23.  
XX New peptide derivatives of heliomicine, useful as antibacterial and  
XX antifungal agents in human or veterinary medicine and agriculture -







```

b      1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrArgAspCysAsnGly 20
Y      61 GAGTGCAGAGAGGGGTTACAAAGGGTGGTCACTGCGGTTCTTCCCTAACGTGAACCTGC 120
b      21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40
Y      121 TGGTGCAGAGACT 132
b      41 TrpCysGluThr 44

RESULT 12
BB76867
D      ABB76867 standard; Protein; 44 AA.
C      ABB76867;
X      03-JUL-2002 (first entry)
X      Antifungal/antibacterial peptide pEM22.
X      Antifungal; antibacterial; fungicide; heliomicine; human medicine;
W      veterinary medicine; plant protection; bacterial infection;
W      fungal infection.
W      Synthetic.
X      WO200206324-A2.
X      24-JAN-2002.
D      05-JUL-2001; 2001WO-FR02164.
X      13-JUL-2000; 2000FR-0009248.
R      19-SEP-2000; 2000FR-0011949.
X      (ENTO-) ENTOMED SA.
X      Dimarcq J, Legrain M, Menin L;
I      WPI; 2002-179779/23.
R      New peptide derivatives of heliomicine, useful as antibacterial and
T      antifungal agents in human or veterinary medicine and agriculture.
S      Claim 15; Page 57; 100pp; French.
X      The present invention relates to peptide derivatives of heliomicine with
C      antifungal and/or antibacterial activity. The present sequence is one
C      such peptide. The peptides are useful in human and veterinary medicine
C      and in plant protection, for control of bacterial and fungal infections.
C      An antifungal and/or antibacterial composition containing at least one
C      peptide is also claimed.
X      Q      Sequence 44 AA;
      Alignment Scores:
      Pred. No.: 2,166-26 Length: 44
      Score: 262.00 Matches: 43
      Percent Similarity: 97.73% Conservative: 0
      Best Local Similarity: 97.73% Mismatches: 1
      Query Match: 98.13% Indels: 0
      DB: 23 Gaps: 0

S-09-673-274B-2_COPY_1_132 (1-132) x ABB76867 (1-44)
Y      1 GATAAGCTTATCGGTTCTCGGTGGGTGCTGTGAACACTACACTTCCGATTGCAACGGT 60
b      1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20
Y      61 GAGTGCAGAGAGGGGTTACAAAGGGTGGTCACTGCGGTTCTTCCCTAACGTGAACCTGC 120
b      21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40

```

```

QY      121 TGGTGCAGACT 132
Db      41 TrpCysGluThr 44

RESULT 13
ABB76869
ID      ABB76869 standard; Protein; 44 AA.
XX      ABB76869;
XX      03-JUL-2002 (first entry)
XX      Antifungal/antibacterial peptide pEM25.
XX      Antifungal; antibacterial; fungicide; heliomicine; human medicine;
KW      veterinary medicine; plant protection; bacterial infection;
KW      fungal infection.
XX      Synthetic.
XX      WO200206324-A2.
XX      24-JAN-2002.
XX      05-JUL-2001; 2001WO-FR02164.
PR      13-JUL-2000; 2000FR-0009248.
PR      19-SEP-2000; 2000FR-0011949.
XX      (ENTO-) ENTOMED SA.
XX      Dimarcq J, Legrain M, Menin L;
DR      WPI; 2002-179779/23.
PT      New peptide derivatives of heliomicine, useful as antibacterial and
PT      antifungal agents in human or veterinary medicine and agriculture.
PS      Claim 15; Page 57; 100pp; French.
XX      The present invention relates to peptide derivatives of heliomicine with
CC      antifungal and/or antibacterial activity. The present sequence is one
CC      such peptide. The peptides are useful in human and veterinary medicine
CC      and in plant protection, for control of bacterial and fungal infections.
CC      An antifungal and/or antibacterial composition containing at least one
CC      peptide is also claimed.
X      Q      Sequence 44 AA;
      Alignment Scores:
      Pred. No.: 2,166-26 Length: 44
      Score: 262.00 Matches: 43
      Percent Similarity: 97.73% Conservative: 0
      Best Local Similarity: 97.73% Mismatches: 1
      Query Match: 98.13% Indels: 0
      DB: 23 Gaps: 0

US-09-673-274B-2_COPY_1_132 (1-132) x ABB76869 (1-44)
QY      1 GATAAGCTTATCGGTTCTCGGTGGGTGCTGTGAACACTACACTTCCGATTGCAACGGT 60
Db      1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20
QY      61 GAGTGCAGAGAGGGGTTACAAAGGGTGGTCACTGCGGTTCTTCCCTAACGTGAACCTGC 120
Db      21 GluCysLysArgArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40
QY      121 TGGTGCAGACT 132
Db      41 TrpCysGluThr 44

RESULT 14
ABB76871

```

10 AB976871 standard; Protein; 44 AA.  
XX  
AC AB976871;  
XX  
XX 03-JUL-2002 (first entry)  
XX  
XX Antifungal/antibacterial peptide p5M7.  
XX  
XX Antifungal; antibacterial; fungicide; heliomicine; human medicine;  
KW veterinary medicine; plant protection; bacterial infection;  
KW fungal infection.  
XX  
XX Synthetic.  
OS  
XX WQ20206324 A2.  
XX  
XX 24 JAN-2002.  
XX  
XX 05 JUL-2001; 2001WO-FR02164.  
XX  
XX 13-JUL-2000; 2000FR-0009248.  
XX  
XX 19 SEP-2000; 2000FR-0011949.  
XX  
XX (ENTO) ; ENTOMED SA.  
XX  
XX D. Mary J. Legrain M. Menu L;  
XX  
XX WPI; 2002 199779/23.  
XX  
XX New peptide derivatives of heliomicine, useful as antibacterial and  
XX antifunga; agents in human or veterinary medicine and agriculture .  
XX  
XX Claim 14; Page 57; 100pp; French.  
XX  
XX The present invention relates to peptide derivatives of heliomicine with  
XX antifungal and/or antibacterial activity. The present sequence is one  
XX such peptide. The peptides are useful in human and veterinary medicine  
XX and in plant protection, for control of bacterial and fungal infections.  
XX An antifungal, and/or antibacterial, composition containing at least one  
XX CC peptide is also claimed.  
XX  
XX  
SQ Sequence 44 AA;  
Alignment Scores:  
Pred. No.: 2,16e-26 Length: 44  
Score: 262.00 Matches: 43  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.13% Indels: 0  
DB: 23 Gaps: 0  
US-09-673-274b-2\_COPY\_1\_132 (1-132) x AB976871 (1-44)  
QY : GATAAGCTTATCGGTCTCGGTGGGTCTGTGAACCTACACTTCGATTGCAACGGT 60  
DB : AspLysLeuIleGlySerCysValTrrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20  
QY 6: GAGTGCAGAGAGGGGTACAGGGTGGTCACTGCGTTCCCTCGGTAAAGTGAACGTCG 120  
DB 2: GluCysLysArgArgGlyTyrLysGlyGlyHicCysGlySerPheAlaAsnValAsnCys 40  
QY 121 TGGTGCAG 129  
DB 41 TrpCysGlu 43  
RESULT 15  
ID AAY54376  
XX AAY54376 standard; Peptide; 74 AA.  
XX  
XX  
XX 06-APR-2000 (first entry)  
XX

DE A fusion peptide of PR-laipha signal peptide/heliomycin.  
XX  
XX PR-laipha; heliomycin; polycysteine peptide; antifungal; antibacterial;  
KW fungal infection; plant protection; animal protection; transgenic plant;  
KW fungi; bacteria; Cercospora beticola; Cladosporium herbarum;  
KW Fusarium culmorum; F. oxysporum; Phytographora cinnamoni.  
XX  
XX Chimeric - Nicotiana sp.  
OS Chimeric - Heliothis virescens.  
XX  
XX Key Location/Qualifiers  
XX Peptide 1..30  
XX Peptide /note= "tobacco PR-laipha signal peptide"  
XX Peptide 31..74  
XX /note= "heliomycin"  
XX  
XX FR2777568-A1.  
XX  
XX 22-OCT-1999.  
XX  
XX 15-APR-1998; 98FR-C004933.  
XX  
XX 15-APR-1998; 98FR-0004933.  
XX  
XX (RHON) RHONE-POULENC AGROCHIMIE.  
XX  
XX WPI; 2000-108532/10.  
XX  
XX N-PSDB; AA245693.  
XX  
XX New polycysteine peptides, designated heliomycin, with antifungal or  
XX antibacterial activity, useful in medicine or plant protection .  
XX  
XX Claim 13; Page 32; 46pp; French.  
XX  
XX The present sequence represents a fusion peptide of the tobacco  
XX PR-laipha signal peptide and heliomycin. The attachment of the  
XX PR-laipha sequence to the N-terminal aids secretion of the heliomyin  
XX peptide. Heliomycin is a polycysteine peptide with antifungal and  
XX antibacterial activity, isolated from the lepidopteron Heliothis  
XX virescens. Heliomycin peptides contain the sequence given in  
XX AAY54378. The heliomyin peptides are used as pharmaceuticals for  
XX treating or preventing fungal infections in humans and animals. They  
XX may also be used as antifungal agents for plant protection. Nucleic  
XX acid encoding the heliomyin peptides is used to generate transgenic  
XX plants that are resistant to some fungi and bacteria (specifically  
XX Cercospora beticola, Cladosporium herbarum, Fusarium culmorum,  
XX F. oxysporum and Phytographora cinnamoni), and for recombinant production  
XX of the peptides.  
XX  
XX Sequence 74 AA;  
SQ  
Alignment Scores:  
Pred. No.: 2,32e-26 Length: 74  
Score: 262.00 Matches: 44  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.13% Indels: 0  
DB: 21 Gaps: 0  
US-09-673-274b-2\_COPY\_1\_132 (1-132) x AAY54376 (1-74)  
QY 1 GATAAGCTTATCGGTCTCGGTGGGTGGTGAACCTACACTTCGATTGCAACGGT 60  
DB 31 AspLysLeuIleGlySerCysValTrrpGlyAlaValAsnTyrThrSerAspCysAsnGly 50  
QY 61 GAGTGCAGAGAGGGGTACAGGGTGGTCACTGCGTTCCCTCGGTAAAGTGAACGTCG 120  
DB 51 GluCysLysArgArgGlyTyrLysGlyGlyHicCysGlySerPheAlaAsnValAsnCys 70  
QY 121 TGGTGCAGACT 132  
DB 71 TrpCysGluThr 74

search completed: October 16, 2003, 17:10:36  
ob time : 41.7841 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

M nucleic - protein search, using frame\_plus\_n2p model

un on: October 16, 2003, 17:06:33 : Search time : 1.2292 Seconds  
(without alignments)  
1415.504 Million cell updates/sec

title: US-09-673-274B-2  
effect score: 324  
sequence: 1 gataagcttatcggttcctg.....gaacgtgcacggatccgg 169

scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

searched: 127863 seqs, 47026705 residues

total number of hits satisfying chosen parameters: 255726

minimum DB seq length: 0  
maximum DB seq length: 2000000000

post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

command line parameters:

MODEL=frame+np.model -DRV=xlp  
Q=/cgn2\_1/USPTO\_epool/US09673274/runat\_16102003\_170627\_2531/app\_query.fasta.1.654  
DB=SwissProt 41 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LST=45  
DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=1.5 -MODE=LOCAL  
OUTFMT=pico -NORM=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000  
USPR=US09673274\_QCGN\_1\_1\_36 -runat\_16102003\_170627\_2531 -NCPU=6 -ICPU=3  
NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
DEV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

result	No.	Score	Query Match	Length	ID	Description
1	267	82.4	44	1	DEFN HELVI	P81544 heliothis v
2	86	26.5	66	1	SCX3_BUTOM	P13488 buthus occi
3	84.5	26.1	85	1	SCN2_MESMA	Q9bX11 mesobuthus
4	84	25.9	85	1	SCAS_MESMA	Q9uac9 mesobuthus
5	83.5	25.8	85	1	SCN3_MESMA	Q9bX10 mesobuthus
6	79.5	24.5	66	1	SCXC_CENLL	P45667 centruoide
7	79	24.4	85	1	SCXE_BUTOC	Q17254 buthus occi
8	78.5	24.2	65	1	SCX4_BUTOM	P59354 buthus occi
9	78.5	24.2	84	1	SCX3_ANDAU	P01480 androctonus
10	77.5	23.9	64	1	SCL6_LEIQH	P59356 leliurus qui
11	77.5	23.9	65	1	SCX4_LEIQH	P01489 leliurus qui
12	77	23.8	85	1	SCAT_MESMA	Q9uac8 mesobuthus
13	76.5	23.6	61	1	SIX2_LEIQH	P19855 leliurus qui
14	76	23.5	66	1	SCXE_BUTEU	P09982 buthus eupe
15	76	23.5	67	1	SCL3_LEIQH	P56678 leliurus qui
16	75.5	23.3	84	1	SCXA_MESMA	O61705 mesobuthus
17	75	23.1	84	1	SCX4_ANDAU	P45658 androctonus
18	74.5	23.0	70	1	DMYC_DRONE	P41964 drosophila

19	74.5	23.0	87	1	SCX1_CENSC	P01492 centruoide
20	74	22.8	59	1	SCX5_CENSC	P58779 centruoide
21	73	22.5	61	1	SIX5_ANDAU	P81504 androctonus
22	73	22.5	64	1	SIX5_LEIQH	P01481 leliurus qui
23	72	22.2	85	1	SC15_MESMA	Q9gng8 mesobuthus
24	72	22.2	85	1	SCA1_MESMA	P80962 buthacus ar
25	71.5	22.1	61	1	SIX2_BUTAR	P55903 buthus occi
26	71.5	22.1	61	1	SIX4_BUTOC	P55904 buthus occi
27	71.5	22.1	61	1	SIX5_BUTOC	P01493 centruoide
28	71.5	22.1	87	1	SCX2_CENSC	P01494 centruoide
29	71.5	22.1	87	1	SCX3_CENSC	P80407 palomera fr
30	71	21.9	43	1	DEF1_PALFR	P59357 leliurus qui
31	71	21.9	66	1	SCL7_LEIQH	P58488 mesobuthus
32	70.5	21.8	64	1	SCM2_MESMA	Q9gng6 mesobuthus
33	70.5	21.8	84	1	SCL6_MESMA	P45664 centruoide
34	70.5	21.8	87	1	SCX6_CENNO	P45697 mesobuthus
35	69.5	21.5	84	1	SCX1_MESVA	P01495 centruoide
36	69.5	21.5	84	1	SCX2_CENNO	Q95wc9 centruoide
37	69.5	21.5	84	1	SCX9_CENSC	Q9n1c5 mesobuthus
38	69.5	21.5	85	1	SC10_MESMA	P82814 buthus sand
39	68.5	21.1	52	1	SIX4_BUTSI	P01488 buthus occi
40	68.5	21.1	65	1	SCX1_BUTOC	P56743 androctonus
41	68.5	21.1	66	1	SCX6_ANDAU	PC9981 buthus eupe
42	68.5	21.1	74	1	SCX9_BUTEU	O65740 capsicum an
43	68.5	21.1	74	1	DEF2_CAPAN	Q9n1c4 mesobuthus
44	68.5	21.1	74	1	SCL7_MESMA	P01479 androctonus
45	68.5	21.1	83	1	SCX1_ANDAU	

ALIGNMENT-S

RESULT 1

ID	DEFN HELVI	STANDARD	PRT	44 AA
AC	P81544;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Defensin heliothisin.			
OS	Heliothis virescens (Noctuid moth) (Owlet moth).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Lepidoptera; Glossata; Noctuoidea;			
OC	Noctuidae; Heliothinae; Heliothis.			
OX	NCBI_TaxID=7102;			
RN	[1]_TaxID=7102;			
RP	SEQUENCE, AND CHARACTERIZATION			
RC	T-SSUE=Hemolymph;			
RX	MEDLINE=99194775; PubMed=1092609;			
RA	Lambert M., Ades S., Uttenweiler-Joseph S., Brookhart G., Bushey D.,			
RA	Hoffmann J.A., Bulet P.			
RT	"Insect immunity. Isolation from the lepidopteran Heliothis virescens			
RT	of a novel insect defensin with potent antifungal activity.";			
RL	J. Biol. Chem. 274:9320-9326 (1999).			
CC	!- FUNCTION: THIS PEPTIDE HAS POTENT ANTI-FUNGAL ACTIVITY. HAS NC			
CC	ACTIVITY AGAINST GRAM-NEGATIVE AND GRAM-POSITIVE BACTERIA			
CC	!- SUBCELLULAR LOCATION: Secreted.			
CC	!- MASS SPECTROMETRY: MW=4784.8; METHOD=MALDI.			
DR	PDB; 1f2v; 12-FEB-02.			
DR	PDB; 1f2v; 12-FEB-02.			
FT	DISULFID 7 32			
FT	DISULFID 18 40			
FT	DISULFID 22 42			
SQ	SEQUENCE 44 AA; 4790 MW; 3CIDA:BD3376FCD5 CRC64;			

Alignment Scores:  
Pred. No.: 3 17e-25 Length: 44  
Score: 267.00 Matches: 44  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 82.41% Indels: 0  
DB: 0 Gaps: 0

US-09-673-274B-2 (1-169) x DEFN\_HELVI (1-44)

QY 1 GATAGCTTATCGTTCTCGGTGGTGGTCTGTGAACCTACACTTCGATTCGACGGT 60  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DB 1 AspLysLeullelGySerCysValrPcYAlaValasnTyrThrSerAspCysasnGly 20  
 QY 61 CAGTCAAGAGAGGGGTTCACAGGGTGGTCACTGCCTTCCTTCCTACGTGAAGTGC 120  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DB 21 GLucLysAlaGAGlyTyTyrLysGlyHisCysGlySerPheA.aAsnValasnCys 40  
 QY 121 TGTCTCGAGACT 132  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DB 41 TrpCysGluThr 44

# RESULT 2

SCX3 BUTOM STANDARD; PRT; 66 AA.

AC P13488;  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Alpha-like neurotoxin II (Bom II) (Bo3).  
 OS Butus occitanus mardochei (Moroccan scorpion).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 OC Butidae; Buthidae; Buthus  
 NCBI\_TaxID=6869;  
 RN 1;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Venom;  
 RA MEDLINE=8161829; PubMed=3104936;  
 RA Vargas O., Martin M.F., Rochat H.;  
 RC "Characterization of six toxins from the venom of the Moroccan  
 scorpion Butus occitanus mardochei";  
 RL Eur J Biochem. 162:589-599(1987).  
 CC 1- FUNCTION: Binds to sodium channels and inhibits the inactivation  
 of the activated channels, thereby blocking neuronal transmission.  
 CC 1 SUBCELLULAR LOCATION: Secreted.  
 CC 1 TISSUE SPECIFICITY: Expressed by the venom gland.  
 CC 1 SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
 CC ALPHA-TOXIN SUBFAMILY. ALPHA-LIKE TOXIN SUBCLASS.

PIR: A29386; A29386.  
 DR HSSP; P56678; LBMR.  
 DR InterPro; IPR003614; Knot1.  
 DR InterPro; IPR002661; Scorpion\_toxinL.  
 DR Pfam; PF00537; toxin\_3; 1.  
 DR ProDom; PD000908; Scorpion\_toxinL; 1.  
 DR SMART; SM00505; Knot1; 1.  
 KW Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.  
 FT DISULFID 12 65  
 FT DISULFID 16 37 BY SIMILARITY.  
 FT DISULFID 23 47 BY SIMILARITY.  
 FT DISULFID 27 49 BY SIMILARITY.  
 SQ SEQUENCE 66 AA; 6872 MW; 7A96DA393F89852F CRC64;

Alignment Scores:  
 Pred. No.: 0.0093 Length: 66  
 Score: 86.00 Matches: 5  
 Percent Similarity: 55.00% Conservative: 7  
 Best Local Similarity: 37.50% Mismatches: 16  
 Query Match: 26.54% Indels: 2  
 DB: 1 Gaps: 1

US-09-673-274B-2 (1-169) x SCX3 BUTOM (1-66)

QY 16 TCTCGGTGGGTGCTGTGAACCTACACTTCGATTCGACGGTGGCAGAGGAGG 75  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DB 11 AsnCysValTyrHisCysPheProGlySerSerGlyCysAspThrLeuCysLysGluLys 30  
 QY 76 GGTACAGGGTGGTCACTCGGTTCCTTC-----GCTAACGTGAAGTGGTGGCGAG 129  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DB 31 GYalaThrSerGlyHisCysGlyPheLeuProGlySerGlyValAlaCysAsp 50

# RESULT 3

## SCN2\_MESMA

ID SCN2\_MESMA STANDARD; PRT; 85 AA.  
 AC Q9BKJ1;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Anti-neuroexcitation peptide II precursor (ANEP11).  
 OS Mesobuthus martensii (Mancurian scorpion) (Buthus martensii).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 OC Buthoidea; Buthidae; Mesobuthus.  
 NCBI\_TaxID=34649;  
 RN 1;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Venom gland;  
 RA Zhang J.-H., Hua Z.C., Zhu D.X.;  
 RT "Cloning of anti-neuroexcitation peptide II (ANEP) cDNA from Scorpion  
 Buthus martensii Karsch";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 CC 1- FUNCTION: Binds to sodium channels and inhibits them.  
 CC 1 SUBCELLULAR LOCATION: Secreted (by similarity).  
 CC 1 TISSUE SPECIFICITY: Expressed by the venom gland.  
 CC 1 SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
 CC ALPHA-TOXIN SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

EMBL; AF242736; AAK28341.1;  
 HSSP; P0491; IB3C.  
 DR InterPro; IPR003614; Knot1.  
 DR InterPro; IPR002661; Scorpion\_toxinL.  
 DR Pfam; PF00537; toxin\_3; 1.  
 DR ProDom; PD000908; Scorpion\_toxinL; 1.  
 DR SMART; SM00505; Knot1; 1.  
 KW Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;  
 KW Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 85 ANTI-NEUROEXCITATION PEPTIDE II.  
 FT DISULFID 3 81 BY SIMILARITY.  
 FT DISULFID 35 56 BY SIMILARITY.  
 FT DISULFID 42 63 BY SIMILARITY.  
 FT DISULFID 46 65 BY SIMILARITY.  
 SQ SEQUENCE 85 AA; 9224 MW; C045C7DACB3513AB CRC64;

Alignment Scores:  
 Pred. No.: 0.014 Length: 85  
 Score: 84.50 Matches: 15  
 Percent Similarity: 55.26% Conservative: 6  
 Best Local Similarity: 39.47% Mismatches: 12  
 Query Match: 26.08% Indels: 5  
 DB: 1 Gaps: 2

US-09-673-274B-2 (1-169) x SCN2\_MESMA (1-85)

QY 16 TCTCGGTGGGTGCTGTGAACCTACACTTCGATTCGACGGTGGCAGAGGAGG 75  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DB 34 SerCysLeuTrpGly-----AsnAspGlyCysAsnLysGluCysArgAlaTyr 49  
 QY 76 GGTACAGGGTGGTCACTCGGTTCCTTCGCTAACGTGAAGTGGTGGCGAG 129  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DB 50 GlyAlaSerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

## RESULT 4

SCAS\_MESMA STANDARD; PRT; 85 AA.  
 ID SCAS\_MESMA  
 AC Q9UAC9; 2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)

```

T 28-FEB-2003 (Rel. 41, Last annotation update)
E Neurotoxin AS precursor.
S Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
C Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
C Buthoidea; Buthidae; Mesobuthus.
C NCBI_TaxID=34649;
N [1]
N SEQUENCE FROM N.A.
P TISSUE=Venom gland;
P TISSUE=Venom gland;
A Lan Z.D., Dai L., Chi C.-W.;
T "The gene cloning and sequencing of AS and AS-1, two novel neurotoxins
T from the scorpion Buthus martensii Karsch.";
L Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
C -!- FUNCTION: Ryanodine receptor agonist in skeletal muscle.
C -!- SUBCELLULAR LOCATION: Secreted (By similarity).
C -!- TISSUE SPECIFICITY: Expressed by the venom gland.
C -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
C ALPHA-TOXIN SUBFAMILY.
C -----
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL Outstation.
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See http://www.isb-sib.ch/announce/
C or send an email to license@isb-sib.ch).
C -----
C EMBL; AF079060; AAD47374.1;
C HSSP; P01492; 1VNB.
C InterPro; IPR003614; Scorpion_toxinL.
C InterPro; IPR002061; Scorpion_toxinL.
C Pfam; PF00537; toxin_3; 1.
C ProDom; PD000908; Scorpion_toxinL; 1.
C SMART; SM00505; Knot1; 1.
C Toxin; Neurotoxin; Signal.
C CHAIN 1 19 POTENTIAL.
C DISULFID 20 85 NEUROTOXIN AS.
C DISULFID 31 81 BY SIMILARITY.
C DISULFID 35 56 BY SIMILARITY.
C DISULFID 42 63 BY SIMILARITY.
C DISULFID 46 65 BY SIMILARITY.
C DISULFID 46 65 BY SIMILARITY.
C SEQUENCE 85 AA; 9759 MW; 41286F9DDCD4FE90 CRC64;
C -----
C Alignment Scores:
C Pred. No.: 0.0161 Length: 85
C Core: 84.00 Matches: 19
C Percent Similarity: 55.81% Conservatives: 5
C Best Local Similarity: 44.19% Mismatches: 17
C Query Match: 25.93% Indels: 2
C B: 1 Gaps: 2
C -----
S-09-673-274B-2 (1-169) x SCAS_MESMA (1-85)
Y 1 GATAAGCTTATCGGTCCTCGGTGGGTGCTGCAACTACACTTCCGATTCGACGCGT 60
||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
b 26 AsplysTyrThrGlyCysLysValTrpCysValIleAsnGluSer---CysAsnSer 44
||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Y 61 GAGTGCAGAGGAGGGGTTACAGAGGTGTCACAGGTGCTTCCTCGCTACAGTGAAGTGC 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
b 45 GluCysLysIleArgGlyGlyTyrTyrCys---TyrPheTrpLysLeuAlaCys 63
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 121 TGGTGGCGAG 129
||| |||||
b 64 PheCysGln 66
||| |||||
C -----
RESULT 5
D_SCN3_MESMA STANDARD; PRT; 85 AA.
C Q9BKJ0; Q9GZC4;
T 28-FEB-2003 (Rel. 41, Created)
T 28-FEB-2003 (Rel. 41, Last sequence update)
T 28-FEB-2003 (Rel. 41, Last annotation update)
DE Anti-neuroexcitation peptide III precursor (ANEPIII) (Anti-epilepsy
peptide).
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Mesobuthus.
OX NCBI_TaxID=34649;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RC TISSUE=Venom gland;
RA Zhang J.-H., Hua Z.C., Zhu D.X.;
RT "Cloning of anti-epilepsy peptide cDNA from scorpion Buthus martensii
RT Karsch.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RC TISSUE=Venom gland;
RA Zhang J.-H., Hua Z.C., Zhu D.X.;
RT "Cloning of anti-neuroexcitation peptide III (ANEPI) cDNA from Scorpion
RT Buthus martensii Karsch.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Binds to sodium channels and inhibits them.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC ALPHA-TOXIN SUBFAMILY.
CC -----
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL Outstation.
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See http://www.isb-sib.ch/announce/
C or send an email to license@isb-sib.ch).
C -----
C EMBL; AF122003; AAG01571.1;
C EMBL; AF242737; AAK28342.1;
C HSSP; P01494; 2SN3.
C InterPro; IPR003614; Knot1.
C InterPro; IPR002061; Scorpion_toxinL.
C Pfam; PF00537; toxin_3; 1.
C ProDom; PD000908; Scorpion_toxinL; 1.
C SMART; SM00505; Knot1; 1.
C Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
C Signal.
C CHAIN 1 21 POTENTIAL.
C DISULFID 22 85 ANTI-NEUROEXCITATION PEPTIDE III.
C DISULFID 31 81 BY SIMILARITY.
C DISULFID 35 56 BY SIMILARITY.
C DISULFID 42 63 BY SIMILARITY.
C DISULFID 46 65 BY SIMILARITY.
C DISULFID 47 47 K -> I (IN REF. 2).
C CONFLICT 47 47
C SEQUENCE 85 AA; 9271 MW; C1D6C93A2F82F8C2 CRC64;
C -----
C Alignment Scores:
C Pred. No.: 0.0186 Length: 85
C Core: 83.50 Matches: 16
C Percent Similarity: 55.26% Conservatives: 5
C Best Local Similarity: 42.11% Mismatches: 12
C Query Match: 25.77% Indels: 5
C DB: 2 Gaps: 2
C -----
US-09-673-274B-2 (1-169) x SCN3_MESMA (1-85)
QY 16 TCCTCGGTGTCGGGTGCTGTGCAACTACACTTCCGATTCGACGCGTGAAGAGGAGG 75
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 34 SerCysLeuTrpGly-----AsnGluGlyCysAsnLysGluCysLysGlyPhe 49
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 76 GGTTCACAGGTTGTCACGTGCGGTTCTCGCTACGTGCAACTGCTGCTGCGAG 129
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 50 GlyA.aTyrTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 6
SCXC_CENLL STANDARD; PRT; 66 AA.
ID_SCXC_CENLL

```



```

RX MEDLINE=9630C228; PubMed=8706664;
RA Bouhaouala-Zahar B., Ducanaci F., Zenouaki I., Ben Khalifa R.,
RA Bourhani L., Peihate M., Boulain J.C., el Ayyeb M., Menez A.,
RA Karoui H.;
RT "A recombinant insect-specific alpha-toxin of Butus occitanus
RT tunetanus scorpion confers protection against homologous mammal
RT toxins.";
RRL Eur. J. Biochem. 238:653-660(1996).
CCC -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CCC of the activated channels, thereby blocking neuronal transmission.
CCC This toxin is active only on insects.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC ALPHA-TOXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X92376; CAA66320.1; -.
DR PIR; S68906; S68906.
DR HSP; P56678; IBMK.
DR InterPro; IPRO03614; KnotI.
DR InterPro; IPRO02061; Scorpion_toxinL.
DR Pfam; PF02537; toxin_3; 1.
DR ProDom; PD000908; Scorpion_toxinL; 1.
DR SMART; SM00505; KnotI; 1.
DR Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW Signal.
KW SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 85 NEUROTOXIN XIV.
FT DISULFID 30 83 BY SIMILARITY.
FT DISULFID 34 55 BY SIMILARITY.
FT DISULFID 41 65 BY SIMILARITY.
FT DISULFID 45 67 BY SIMILARITY.
FT SEQUENCE 85 AA; 9174 MW; 4977F16C511F0567 CRC64;
SQ
Alignment Scores:
Pred. No.: 0.0671 Length: 85
Score: 79.00 Matches: 14
Percent Similarity: 50.00% Conservative: 6
Best Local Similarity: 35.00% Mismatches: 18
Query Match: 24.38% Indels: 2
DB: 1 Gaps: 1
US-09-673-274B-2 (1-169) x SCXE_BUTOC (1-85)
QY 16 TCCTCGGTGGGTTGTGAACATCACTCCGATTGCACAGCGTGAGTCGAAGAGGAGG 75
Db 29 AsnCysAlaTyHisCysLeuLysIleSerSerGlyCysAspThrLeuCylsGluasn 48
QY 76 GOTTACAGGCGTGCTACTCGGT-----TCCTTCGCTAACGTGAACGTGTCGCGAG 129
Db 49 GYAlaThrSerGlyHisCysGlyHisLysSerGlyHisGlySerAlaCysTrpCyslys 68
RESULT 8
SCX4 BUTOM STANDARD; PRt; 65 AA.
ID _SCX4_BUTOM
AC P59354;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alpha-like neurotoxin IV (Bom IV) (Bom4).
OS Butus occitanus mardochei (Moroccan scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Scorpiones;
OC Buthoidea; Buthidae; Buthus.
RN NCBI_Taxid=6869;
RN [1]

```

```

P SEQUENCE, PHARMACOLOGICAL CHARACTERIZATION, AND MASS SPECTROMETRY.
C TISSUE=Venom;
X MEDLINE=99203324; PubMed=10103091;
A Cestele S., Stankiewicz M., Mansuelle P., De Waard M., Dargent B.,
A Gilles N., Pelhate M., Rochat H., Martin-Eauclaire M.-F., Gordon D.;
T "Scorpion alpha-like toxins, toxic to both mammals and insects,
T differentially interact with receptor site 3 on voltage-gated sodium
T channels in mammals and insects";
L Eur. J. Neurosci. 11:975-985(1999).
C -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
C of the activated channels, thereby blocking neuronal transmission.
C -!- SUBCELLULAR LOCATION: Secreted.
C -!- TISSUE SPECIFICITY: Expressed by the venom gland.
C -!- MASS SPECTROMETRY: MW:7287.96; MW ERR:0.37; METHOD=Electrospray.
C -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
C ALPHA-TOXIN SUBFAMILY. ALPHA-LIKE TOXIN SUBCLASS.
R Pfam: PF00537; toxin_3; 1.
R ProDom: PD000908; Scorpion_toxinL; 1.
R SMART: SM00505; Knt1; 1.
R Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
W DISULFID 12 63 BY SIMILARITY.
T DISULFID 16 36 BY SIMILARITY.
T DISULFID 22 46 BY SIMILARITY.
T DISULFID 26 48 BY SIMILARITY.
Q SEQUENCE 65 AA; 7296 MW; 4305EB32436A4B90 CRC64;

Alignment Scores:
red. No.: 0.0759 Length: 65
core: 78.50 Matches: 16
Percent Similarity: 51.22% Conservative: 5
Best Local Similarity: 39.02% Mismatches: 15
Query Match: 24.23% Indels: 5
B: 1 Gaps: 3

S-09-673-274B-2 (1-169) x SCX4_BUTOM (1-65)
Y 16 TCCTGCGTGTGGGGTGTGTGCAACTACACTTCCGATTGCAACGGTGTGAGTCCAGAGGAGG 75
b :|||:|||||:
b 11 AsnCysValTyrGluCysAlaLys---AaSerTyrCysAsnAspLeuCysThrLysAsn 29
Y 76 GGTTCACAGGGTGTCTACTGC-----GGTTCCTTCCTAACGTGAACCTCTGTGTC 126
b :|||:|||||:|||||:
b 30 GlyAlaLysSerGlyTyrCysGlnTrpLeuGlyLysTyrGlyAsnAla---CysTrpCys 48
Y 127 GAG 129
b :|||
b 49 Glu 49

RESULT 9
CX3 ANDAU STANDARD; PRT; 84 AA.
C P01480;
T 21-JUL-1986 (Rel. 01, Created)
T 01-JUL-1993 (Rel. 26, Last sequence update)
T 28-FEB-2003 (Rel. 41, Last annotation update)
E Neurotoxin III precursor (AaH III) (AaH3).
S Androctonus australis hector (Sahara scorpion).
C Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
C Buthoidea; Buthidae; Androctonus.
X NCBI_taxid=70175;
N [1]
N SEQUENCE FROM N.A.
P MEDLINE=90037062; PubMed=2808423;
X Bougis P.E., Rochat H., Smith L.A.;
A "Precursors of Androctonus australis scorpion neurotoxins. Structures
T of precursors, processing outcomes, and expression of a functional
T recombinant toxin II.";
L J. Biol. Chem. 264:19259-19265(1989).
L [2]
P SEQUENCE OF 20-83.
X MEDLINE=79148414; PubMed=428402;
X Kopeyan C., Martinez G., Rochat H.;
T "Amino acid sequence of neurotoxin III of the scorpion Androctonus

```

```

RT australis Hector. ";
RL Eur. J. Biochem. 94:609-615(1979).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=9112921; PubMed=2279844;
RA Laplante S.R., Mikou A., Robin M., Guittet E., Delsuc M.A.,
RA Charpentier I., Lallemand J.-Y.;
RT "Rapid determination and NMR assignments of antiparallel sheets and
RT helices of a scorpion and a cobra toxin.";
RL Int. J. Pept. Protein Res. 36:227-230(1990).
RN [4]
RP STRUCTURE BY NMR.
RX MEDLINE=93044377; PubMed=1422146;
RA Mikou A., Laplante S.R., Guittet E., Lallemand J.-Y.,
RA Martin-Eauclaire M.-F., Rochat H.;
RT "Toxin III of the scorpion Androctonus australis hector: proton
RT nuclear magnetic resonance assignments and secondary structure.";
RL J. Biomol. NMR 2:57-70(1992).
CC -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC of the activated channels, thereby blocking neuronal transmission.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC ALPHA-TOXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to: license@sib-sib.ch)
CC -----
CR EMBL: M27003; AAA29948.1;
DR PIR: C34444; NTSR3A.
DR HSSP: P56678; IEMR.
DR InterPro: IPR003614; Knt1.
DR InterPro: IPR002961; Scorpion_toxinL.
DR Pfam: PF00537; toxin_3; 1.
DR ProDom: PD000908; Scorpion_toxinL; 1.
DR SMART: SM00505; Knt1; 1.
KW Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW Signal.
FT SIGNAL 1 19
FT CHAIN 20 83 NEUROTOXIN III.
FT PROPEP 84 84
FT DISULFID 31 91 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 39 63 BY SIMILARITY.
FT DISULFID 43 65 BY SIMILARITY.
FT CONFLICT 27 27 D->N (IN REF. 2).
SQ SEQUENCE 84 AA; 9068 MW; ADB0AD462153E3CC CRC64;

Alignment Scores:
Pred. No.: 0.0773 Length: 84
Score: 78.50 Matches: 15
Percent Similarity: 51.85% Conservative: 6
Best Local Similarity: 39.46% Mismatches: 13
Query Match: 24.23% Indels: 5
DB: 1 Gaps: 2

US-09-673-274B-2 (1-169) x SCX3_ANDAU (1-94)
QY 16 TCCTGCGTGTGGGGTGTGTGCAACTACACTTCCGATTGCAACGGTGTGAGTCCAGAGGAGG 75
Db :|||:|||||:|||||:
Db 30 AsnCysValTyrHisCysVal-----ProCysAspGlyLeuCysLysAsn 48
QY 76 GGTTCACAGGGTGTCTACTGCCTTC-----GCTACGTGAACCTCTGTGTC 126
Db :|||:|||||:|||||:
Db 47 GlyAlaLysSerG-ySerCysGlyPheLeuLeuProSerGlyLeuAlaCysTrpCys 65

RESULT 10
SCL6_LEIQH

```

```

RA Kopeyan C., Martinez G., Rochat H.;
RT "Primary structure of toxin IV of Leiurus quinquestriatus
RL quinquestriatus";
RL FEBS Lett. 181:211-217(1985).
CC - FUNCTION. Binds to sodium channels and inhibits the inactivation
CC of the activated channels, thereby blocking neuronal transmission.
CC - SUBCELLULAR LOCATION. Secreted.
CC - TISSUE SPECIFICITY. Expressed by the venom gland.
CC - SIMILARITY. BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC ALPHA-TOXIN SUBFAMILY.
DR PIR; A01749; NTSK4L.
DR HSSP; P45697; ISN1.
DR InterPro; IPR033614; Knot1.
DR InterPro; IPR031219; Neurotoxin.
DR InterPro; IPR032061; Scorpion_toxinL.
DR Pfam; PF05337; toxin_3; 1.
DR PRINTS; PR00284; TOXIN.
DR ProDom; PD000938; Scorpion_toxinL; 1.
DR SMART; SM00505; Knot1; 1.
DR Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW Amidation.
KW DISULFID 13 64 BY SIMILARITY.
FT DISULFID 17 37 BY SIMILARITY.
FT DISULFID 23 47 BY SIMILARITY.
FT DISULFID 27 49 BY SIMILARITY.
FT MOD_RES 65 65 AMIDATION.
FT MOD_RES 65 65
SQ SEQUENCE 65 AA; 7202 MW; 892102EC1C2B89F6 CRC64;

Alignment Scores:
Pred. No.: 0.01 Length: 65
Score: 77.50 Matches: 16
Percent Similarity: 57.14% Conservative: 8
Best Local Similarity: 38.10% Mismatches: 9
Query Match: 23.92% Indels: 9
DB: 1 Gaps: 4

US-09-673-274B-2 (1-159) x SCX4_LE1QU (1-65)
QY 16 TCCTCGGTGG-----GGTGTGTAATACACTTCGATTGCAACGGTGAGTGCAG 69
Db 12 AsnCysVa.TyrThrCysGlySerAsnSerTyr-----CysAsnThrGluCysThr 28
QY 70 AGAGGGGTTCACAGGGGTGCCTGC-----GGTCTCTTCGTACGCTGAACCTGC 120
Db 29 LysAsnGlyA.GluSerGlyTyrCysGlnTrpLeuGlyLysTyrGlyAsnAla---Cys 47
QY 121 TGGTGC 126
Db 48 TrpCys 49

RESULT 12
SCAT_MESMA
ID SCAT_MESMA STANDARD; PRT; 85 AA.
AC Q9UAC8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurotoxin AS-1 precursor.
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Mesobuthus.
OX NCBI_TaxID=34649;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Lan 2.D.; Dai L.; Chi C.-W.;
RT "The gene cloning and sequencing of AS and AS-1, two novel neurotoxins
RL from the scorpion Buthus martensii Karsch.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION. Ryanodine receptor agonist in skeletal muscle.
CC - SUBCELLULAR LOCATION. Secreted (by similarity).
CC - TISSUE SPECIFICITY. Expressed by the venom gland.
CC - SIMILARITY. BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.

```





Result No.	Score	Match	Query Length	DB Length	ID	Description
1	84.5	26.1	65	5	Q17231	O17231 mesobuthus

```
DR InterPro: IPR001219; Neurotoxin.
DR InterPro: IPR002061; Scorpion_toxinL.
DR Pfam: PF00537; toxin_3; 1.
DR PRINTS; PR02284; TOXIN.
DR PRODOM; PD000908; Scorpion_toxinL; 1.
DR SMART; SM00505; Knot1; 1.
SQ SEQUENCE 65 AA; 7285 MW; F13FCD153855E26 CRC64;

Alignment Scores:
Pred. No.: 6,0313 Length: 65
Score: 84.50 Matches: 15
Percent Similarity: 55.26% Conservative: 6
Best Local Similarity: 39.47% Mismatches: 12
Query Match: 26.08% Indels: 5
DB: Gaps: 2

US-09-673-274B-2 (1-169) x Q17231 (1-65)

QY 16 TCCTCGGTGGGTGCTGTGAACACTCCGATTGCAACGGTGAGTGCAGAGGAGG 75
Db 14 SerCysLeuTrpGly-----AsnGluGlyCysAsn:ysGluCysArgAlaTyr 29
QY 76 GGTACAAAGGTGGTCACTCGGTTCCTTCGCTAACCTGAACCTGCTGGTCCGAG 129
Db 30 GlyAlaTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 46

RESULT 2
Q17230
ID Q17230 PRELIMINARY; PRT; 65 AA.
AC Q17230;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Insect toxin 3.
GN "Molecular cloning and functional expression of a gene encoding an
antiarrhythmia peptide derived from the scorpion toxin.";
RL Eur. J. Biochem. 269:4468-4475(2002).
DR EMBL; AF459791; AAN59782.1; .
DR EMBL; AF459792; AAN59783.1; .
DR EMBL; AF459793; AAN59784.1; .
SQ SEQUENCE 65 AA; 7317 MW; ECDCC391B733EC0A CRC64;

Alignment Scores:
Pred. No.: 0.0422 Length: 65
Score: 83.50 Matches: 16
Percent Similarity: 55.26% Conservative: 5
Best Local Similarity: 42.11% Mismatches: 12
Query Match: 25.77% Indels: 5
DB: Gaps: 2

US-09-673-274B-2 (1-169) x Q17230 (1-65)

QY 16 TCCTCGGTGGGTGCTGTGAACACTCCGATTGCAACGGTGAGTGCAGAGGAGG 75
Db 14 SerCysLeuTrpGly-----AsnGluGlyCysAsn:ysGluCysLysGlyPhe 29
```

```
QY 76 GGTACAAAGGTGGTCACTCGGTTCCTTCGCTAACCTGAACCTGCTGGTCCGAG 129
Db 30 GlyAlaTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 46

RESULT 3
Q810K7
ID Q810K7 PRELIMINARY; PRT; 85 AA.
AC Q810K7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Depressant scorpion toxin KIM2 precursor.
GN KIM2.
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Mesobuthus.
OX NCBI_TaxID=34649;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=22217688; PubMed=12230558;
RA Peng F., Zeng X.C., He X.H., Pu J., Li W.X., Zhu Z.H., Liu H.;
RT "Molecular cloning and functional expression of a gene encoding an
antiarrhythmia peptide derived from the scorpion toxin.";
RL Eur. J. Biochem. 269:4468-4475(2002).
DR EMBL; AF459791; AAN59782.1; .
DR EMBL; AF459792; AAN59783.1; .
DR EMBL; AF459793; AAN59784.1; .
SQ SEQUENCE 85 AA; 9425 MW; 5076039B966AA6F CRC64;

Alignment Scores:
Pred. No.: 0.043 Length: 85
Score: 83.50 Matches: 16
Percent Similarity: 55.26% Conservative: 5
Best Local Similarity: 42.11% Mismatches: 12
Query Match: 25.77% Indels: 5
DB: Gaps: 2

US-09-673-274B-2 (1-169) x Q810K7 (1-85)

QY 16 TCCTCGGTGGGTGCTGTGAACACTCCGATTGCAACGGTGAGTGCAGAGGAGG 75
Db 34 SerCysLeuTrpGly-----AsnGluGlyCysAsnLysGluCysLysGlyPhe 49
QY 76 GGTACAAAGGTGGTCACTCGGTTCCTTCGCTAACCTGAACCTGCTGGTCCGAG 129
Db 50 GlyAlaTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66

RESULT 4
Q95WX6
ID Q95WX6 PRELIMINARY; PRT; 85 AA.
AC Q95WX6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Depressant insect neurotoxin Bmk ITb precursor.
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Mesobuthus.
OX NCBI_TaxID=34649;
RN [1];
RP SEQUENCE FROM N.A.
RA Wang C., Ling M., Wang D., Chi C.;
RT "The gene for depressant insect toxin Bmk ITb from Chinese scorpion,
Buthus martensii Karsch.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
DR EMBL; AF272777; AAF77063.2; .
DR InterPro: IPR003614; Knot1.
DR InterPro: IPR002061; Scorpion_toxinL.
DR Pfam; PF00537; toxin_3; 1.
DR PRODOM; PD000908; Scorpion_toxinL; 1.
DR SMART; SM00505; Knot1; 1.
DR Neurotoxin; Signal.
```

T SIGNAL 1 21 POTENTIAL.  
 T CHAIN 22 82  
 Q SEQUENCE 85 AA; 9330 MW; 80CDD3914956DB8C CRC64;  
 Alignment Scores:  
 Pred. No.: 0.0783 Length: 85  
 Score: 81.50 Matches: 15  
 Percent Similarity: 55.26% Conservative: 6  
 Best Local Similarity: 39.47% Mismatches: 12  
 Query Match: 25.15% Indels: 5  
 DB: 5 Gaps: 2

US-09-673-274B-2 (1-169) x Q95WX6 (1-85)  
 16 TCCTGGGTGGGGTGGTGGTGAACATACACTTCCGATTCGACAGCGTGAGTGCACAGAGG 75  
 |||||:|||||  
 34 SerCysLeuTrpGly-----AsnGluGlyCysAsnLysGluCysLysAlaPhe 49  
 |||||:|||||  
 76 GGTTACAAAGGTGGTCACTCGGTTTCCTTCCTAACGTAACGTAACGTCGTCGTCGAG 129  
 |||||:|||||  
 50 GlyAlaSerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGln 66  
 |||||:|||||

## RESULT 5

Q9XY87 PRELIMINARY; PRT; 85 AA.  
 Q9XY87  
 01-NOV-1999 (TrEMBLrel. 12, Created)  
 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Neurotoxin AEP precursor.  
 Mesobuthus martensii (Machurian scorpion) (Buthus martensii).  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 Buthoidea; Buthidae; Mesobuthus.  
 NCBI\_TaxID=34649;  
 [1]  
 SEQUENCE FROM N.A.  
 Ling M.H., Wang C.G., Wang D.C., Chi C.W.;  
 Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).  
 -1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
 ENBL; AF064821; A0031592.1; -;  
 HSP; P01494; 2SN3.  
 InterPro; IPR003614; Knot1.  
 InterPro; IPR002061; Scorpion\_toxinL.  
 Pfam; PF00537; toxin\_3; 1.  
 ProDom; PD000908; Scorpion\_toxinL; 1.  
 SMART; SMO0505; Knot1; 1.  
 Neurotoxin; Signal.  
 T SIGNAL 1 21 POTENTIAL.  
 T CHAIN 22 85 NEUROTOXIN AEP.  
 Q SEQUENCE 85 AA; 9312 MW; CCE46711BAF21DAC CRC64;

Alignment Scores:  
 Pred. No.: 0.0783 Length: 85  
 Score: 81.50 Matches: 14  
 Percent Similarity: 55.26% Conservative: 7  
 Best Local Similarity: 36.84% Mismatches: 12  
 Query Match: 25.15% Indels: 5  
 DB: 5 Gaps: 2

US-09-673-274B-2 (1-169) x Q9XY87 (1-85)  
 16 TCCTGGGTGGGGTGGTGGTGAACATACACTTCCGATTCGACAGCGTGAGTGCACAGAGG 75  
 |||||:|||||  
 34 SerCysLeuTrpGly-----AsnGluGlyCysAsnLysGluCysArgAlaTyr 49  
 |||||:|||||  
 76 GGTTACAAAGGTGGTCACTCGGTTTCCTTCCTAACGTAACGTCGTCGTCGAG 129  
 |||||:|||||  
 50 GlyAlaSerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGln 66  
 |||||:|||||

## RESULT 6

Q8T3T0 PRELIMINARY; PRT; 85 AA.  
 Q8T3T0  
 01-JUN-2002 (TrEMBLrel. 21, Created)  
 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Depressant insect toxin Bmk 1Tal precursor.  
 Mesobuthus martensii (Machurian scorpion) (Buthus martensii).  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 Buthoidea; Buthidae; Mesobuthus.  
 NCBI\_TaxID=34649;  
 [1]  
 SEQUENCE FROM N.A.  
 Liu Z., Chi C., Wu X.;  
 Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).  
 -1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
 ENBL; AY090782; A0030997.1; -;  
 InterPro; IPR003614; Knot1.  
 InterPro; IPR002061; Scorpion\_toxinL.  
 Pfam; PF00537; toxin\_3; 1.  
 ProDom; PD000908; Scorpion\_toxinL; 1.  
 SMART; SMO0505; Knot1; 1.  
 Neurotoxin; Signal.  
 T SIGNAL 1 21 POTENTIAL.  
 T CHAIN 22 85 NEUROTOXIN AEP.  
 Q SEQUENCE 85 AA; 9312 MW; CCE46711BAF21DAC CRC64;

AC Q8T3T0:  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Depressant insect toxin Bmk 1Tal precursor.  
 Mesobuthus martensii (Machurian scorpion) (Buthus martensii).  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 Buthoidea; Buthidae; Mesobuthus.  
 NCBI\_TaxID=34649;  
 [1]  
 SEQUENCE FROM N.A.  
 Liu Z., Chi C., Wu X.;  
 Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).  
 -1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
 ENBL; AY090782; A0030997.1; -;  
 DR InterPro; IPR003614; Knot1.  
 DR InterPro; IPR002061; Scorpion\_toxinL.  
 DR Pfam; PF00537; toxin\_3; 1.  
 DR ProDom; PD000908; Scorpion\_toxinL; 1.  
 DR SMART; SMO0505; Knot1; 1.  
 KW Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 SQ SEQUENCE 85 AA; 9214 MW; 60F5670B6C32001A CRC64;  
 Alignment Scores:  
 Pred. No.: 0.106 Length: 85  
 Score: 80.50 Matches: 15  
 Percent Similarity: 52.63% Conservative: 5  
 Best Local Similarity: 39.47% Mismatches: 13  
 Query Match: 24.85% Indels: 5  
 DB: 5 Gaps: 2

US-09-673-274B-2 (1-169) x Q8T3T0 (1-85)  
 QY 16 TCCTGGGTGGGGTGGTGGTGAACATACACTTCCGATTCGACAGCGTGAGTGCACAGAGG 75  
 |||||:|||||  
 DB 34 SerCysLeuTrpGly-----AsnGluGlyCysAsnLysGluCysGlyAlaTyr 49  
 |||||:|||||  
 QY 76 GGTTACAAAGGTGGTCACTCGGTTTCCTTCCTAACGTAACGTCGTCGTCGAG 129  
 |||||:|||||  
 DB 50 GlyAlaSerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 66  
 |||||:|||||

RESULT 7  
 P82761 PRELIMINARY; PRT; 91 AA.  
 ID P82761  
 AC P82761  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein LCR46 precursor.  
 GN LCR46.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Caryophyllales; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxID=3702;  
 [1]  
 SEQUENCE FROM N.A.  
 RP STRAIN=cv. Columbia;  
 RX MEDLINE=99397451; PubMed=10470850;  
 RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,  
 RA Miyajima N., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.  
 RT Sequence features of the regions of 1,011,550 bp covered by seventeen  
 RT P1 and TAC clones."  
 RL DNA Res. 6:183-195(1999).  
 [2]  
 RN IDENTIFICATION.  
 RP PubMed=1437247;  
 RX Vancosthuyse V., Mieg C., Dumas C., Cock J.M.;  
 RA "Two large Arabidopsis thaliana gene families are homologous to the  
 RT Brassica gene superfamily that encodes pollen coat proteins and the



RT male component of the self-incompatibility response";  
 RL Plant Mol. Biol. 46:17-34(2002).  
 CR EMBL: AB017061; NOT ANNOTATED\_CDS.  
 KW Hypothetical protein; Signal.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 91 HYPOTHETICAL PROTEIN LCR46.  
 SQ SEQUENCE 91 AA; 9503 MW; 130DEAABE8C969F55 CRC64;

## Alignment Scores:

Pred. No.: 0.123 Length: 91  
 Score: 80.00 Matches: 14  
 Percent Similarity: 58.06% Conservative: 4  
 Best Local Similarity: 45.16% Mismatches: 13  
 Query Match: 24.69% Indels: 0  
 DB: 10 Gaps: 0

US-09-673-274B-2 (1-169) x P82762 (1-91)

QY 16 TCCTCGCTGGGCTGCTGTAACACTACACTTCGATTCGAACGCTGAGTGCAGAGGAGG 75

DD 31 ThrCysLeuProGlyCysThrAsnProSerGluCysAsnAlaAlaCysLysSerAsn 50

QY 76 GGTTCACAAAGGTGGTCACTCGGCTTCCTTCGCT 108

DD 5: GlyTyrLysGlyGlyAlaCysValSerSer 61

## RESULT 8

Q9VZQ5 PRELIMINARY; PRT; 69 AA.

AC Q9VZQ5;

LT 01-MAY 2000 (TrEMBLrel. 13, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR 2003 (TrEMBLrel. 23, Last annotation update)

DE G32274 protein.

GN G32268 OR CG10815 OR CG11520.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

CC Ephydriidae; Drosophilidae; Drosophila.

OX NCBI:TaxID=7227;

RN 1:

RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;

RX MEDLINE:20196066; PubMed:10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amaral-Des P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman G.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.J., Andrews-Flamkuch C., Baldwin D.,

RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkeva D., Botchan M.R., Bouck C., Brockstein P., Brothier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Chery J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,

RA dePablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser M.,

RA Giedek A., Gong F., Gortelli J.H., Gu Z., Guo P., Harris K.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li S., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Murray D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Sirpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao O., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).

RN 12:

RP SEQUENCE FROM N.A.

RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,

RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,

RA Bazon J., An H., Baldwin D., Banzon C., Beeson K.Y., Busam D.A.,

RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,

RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,

RA Ferreria S., Frise E., Galle R.F., Garg N.S., George R.A.,

RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,

RA Ibegwam C., Jalali M., Kruse D., Li P., Mattel B., Moshrefi A.,

RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,

RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,

RA Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,

RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,

RA Williams S.M., Zaveri J.S., Smith H.O., Venter C., Rubin G.M.,

RT "Sequencing of Drosophila melanogaster genome";

RL Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.

RN 13:

RP SEQUENCE FROM N.A.

RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,

RA Hradecky P., Huang Y., Kaminker J.S., Prochink S.B., Smith C.D.,

RA Tupy J., Bergman C., Berman B., Carlson J.W., Celniker S.E.,

RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,

RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,

RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,

RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,

RT "Annotation of Drosophila melanogaster genome";

RL Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.

RN 14:

RP SEQUENCE FROM N.A.

RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.,

RL Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.

RN 15:

RP SEQUENCE FROM N.A.

RA FlyBase;

RL Submitted (SEP-2002) to the EMBL/GenBank/DBSJ databases.

DR EMBL: AE003477; AAF47764.2;

DR HSSP: P41964; IMYN.

DR FlyBase: FBgn0052268; CG32268.

DR InterPro: IPR002118; Gamma-thionin.

DR ProDom: PD002594; Gamma-thionin; 1.

SQ SEQUENCE 69 AA; 7720 MW; 87CC05D254A1AC4 CRC64;

## Alignment Scores:

Pred. No.: 0.19 Length: 69  
 Score: 78.50 Matches: 16  
 Percent Similarity: 51.28% Conservative: 4  
 Best Local Similarity: 41.03% Mismatches: 14  
 Query Match: 24.23% Indels: 5  
 DB: 5 Gaps: 2

US-09-673-274B-2 (1-169) x Q9VZQ5 (1-69)

QY 13 GGTTCCTCGTGGGCTGCTGTAACACTACACTTCGATTCGAACGCTGAGTGCAGAGG 72

DB 34 GlySerCys-----AlaValTrpHisArgLysLysCysValaApIleCysGlnArg 50

QY 73 AGGGGTTACAAGGCTGCTACTCGGCTTCCTTCGCTAACGTAACGCTGCTGGAG 129

DB 51 GluGlyArgThrSerGlyHisCysSer-----ProSerLeuLysCysTrpCysGlu 67

## RESULT 9

Q8MV57

ID Q8MV57

PRELIMINARY; PRT; 82 AA.



Qy 34 .....CAGCACCC 26  
Db 73 PheSerProLeuGluHisPro 79

RESULT 12  
Q9VZR2 Q9VZR2 PRELIMINARY; PRT; 69 AA.

ID Q9VZR2 AC Q9VZR2  
01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE CG10812 protein (GH09576p).  
CN CG10812

CS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkely.  
RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.I.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
RA Balaw R.M., Basu A., Bakendale C., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berham B.P., Bhargava P., Brodtier P.,  
RA Boikova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Patris B., Delcher A., Deng Z., Mays A.D., Dew J., Dietz S.M.,  
RA Dodson K.J., Doup L.S., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Feirrer W.M., Glasser K.,  
RA Foslter C., Gabrielson A.E., Gang N.S., Gelbart W.M., Glasner K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jaiall M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp O., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Jiang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson J.,  
RA Merkulov G., Malshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman K.S., Pan S., Pellard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheef F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swislockas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.W., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L.,  
RA Zheng X.H., Zhong F., Zhong W., Zhou X., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter C.C.,  
RA "The genome sequence of Drosophila melanogaster".  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkely.  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
RA Maranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Ceiniher S.;  
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
RL EMBL: AE003477; AAF4757.1;  
DR EMBL: AY118762; AAM50622.1;  
DR HSSP: P41964; IMYN.

Qy 34 .....CAGCACCC 26  
Db 73 PheSerProLeuGluHisPro 79

RESULT 12  
Q9VZR2 Q9VZR2 PRELIMINARY; PRT; 69 AA.

ID Q9VZR2 AC Q9VZR2  
01-MAY-2000 (TReMBLrel\_13, Created)  
DT 01-MAY-2000 (TReMBLrel\_13, Last sequence update);  
DT 01-MAR-2003 (TReMBLrel\_23, Last annotation update)  
DE CG10812 protein (GH09576p).  
CN CG10812

CS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
NC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
CX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkely;  
RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.I.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
RA Balaw R.M., Basu A., Bakendale C., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berlan B.P., Bhargava P., Brodtier P.,  
RA Boikova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,  
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Patris B., Delcher A., Deng Z., Mays A.D., Dew J., Dietz S.M.,  
RA Dodson K.J., Doup L.S., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fodor C., Gabrielson A.E., Gang N.S., Gelbart W.M., Glasser K.,  
RA Goddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jaiall M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp O., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Jiang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson J.,  
RA Merkulov G., Malshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman K.S., Pan S., Pellard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheef F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swislock R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.W., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L.,  
RA Zheng X.H., Zhong F., Zhong W., Zhou X., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter C.C.,  
RA "The genome sequence of Drosophila melanogaster";  
Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkely;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
RA Maranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Ceiniher S.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
RE EMBL: AE003477; AAF4757.1;  
DR EMBL: AY118762; AAM50622.1;  
DR HSSP: P41964; IMYN.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

MM nucleic - protein search, using frame\_plus\_n2p model

un on: October 16, 2003, 17:06:33 Search time 52.2159 Seconds  
(without alignments)  
1027.456 Million cell updates/sec

US-09-673-274b-2

effect score: 324

equences: 1 gataagcttattcgttctctg.....gaacgtgtcagcgatccgg 169

scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODEL=frame\_n2p.model -DEV=xlp

Q=/cgn2\_1/USPRO\_pool/US09673274/runat\_16102003\_170626\_2523/app\_query\_fasta\_1.654

DB=A Geneseq\_19Jun03 -QMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPL=0

LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

LIST=45 -DOCALIGN=200 -THR SCORE=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=15

MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000

USER=US09673274\_QCGN\_1\_135@runat\_16102003\_170626\_2523 -NCPJ=3

NO MMAP -LARGESCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq\_19Jun03.\*

- 1: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSI/cgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSI/cgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSI/cgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDSI/cgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDSI/cgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	267	82.4	44	21	AA54375	Amino acid sequenc
2	267	82.4	44	23	AB76860	Antifungal/antibac
3	267	82.4	44	23	AB76868	Antifungal/antibac
4	267	82.4	49	21	AA54374	Amino acid sequenc
5	267	82.4	74	21	AA54376	A fusion peptide C
6	266	82.1	44	23	AB76895	Antifungal/antibac
7	264	81.5	44	23	AB76863	Antifungal/antibac
8	264	81.5	44	23	AB76864	Antifungal/antibac
9	264	81.5	44	23	AB76896	Antifungal/antibac
10	262	80.9	44	23	AB76862	Antifungal/antibac
11	262	80.9	44	23	AB76865	Antifungal/antibac
12	262	80.9	44	23	AB76867	Antifungal/antibac
13	262	80.9	44	23	AB76869	Antifungal/antibac
14	262	80.9	44	23	AB76871	Antifungal/antibac
15	262	80.9	44	23	AB76871	Antifungal/antibac
16	261	80.6	44	23	AB76866	Antifungal/antibac
17	261	80.6	44	23	AB76873	Antifungal/antibac
18	261	80.6	44	23	AB76894	Antifungal/antibac
19	259	79.9	44	23	AB76873	Antifungal/antibac
20	259	79.6	44	23	AB76898	Antifungal/antibac
21	256	79.0	44	23	AB76861	Antifungal/antibac
22	256	79.0	44	23	AB76897	Antifungal/antibac
23	255	78.7	44	23	AB76882	Antifungal/antibac
24	254	78.4	44	23	AB76874	Antifungal/antibac
25	254	78.4	44	23	AB76875	Antifungal/antibac
26	253	78.1	44	23	AB76879	Antifungal/antibac
27	253	78.1	44	23	AB76885	Antifungal/antibac
28	251	77.5	44	23	AB76876	Antifungal/antibac
29	251	77.5	44	23	AB76877	Antifungal/antibac
30	251	77.5	44	23	AB76881	Antifungal/antibac
31	251	77.5	44	23	AB76884	Antifungal/antibac
32	251	77.5	44	23	AB76886	Antifungal/antibac
33	250	77.2	44	23	AB76880	Antifungal/antibac
34	250	77.2	44	23	AB76883	Antifungal/antibac
35	248	76.5	44	23	AB76887	Antifungal/antibac
36	248	76.5	44	23	AB76888	Antifungal/antibac
37	246	75.9	44	23	AB76878	Antifungal/antibac
38	246	75.9	44	23	AB76889	Antifungal/antibac
39	246	75.9	44	23	AB76890	Antifungal/antibac
40	243	75.0	44	23	AB76892	Antifungal/antibac
41	243	75.0	44	23	AB76899	Antifungal/antibac
42	241	74.4	44	23	AB76891	Antifungal/antibac
43	238	73.5	44	23	AB76893	Antifungal/antibac
44	232	71.6	44	24	ABP55988	Agrotis :psillon Fu
45	232	71.6	73	24	ABP55987	Agrotis :psillon Fu

ALIGNMENTS

RESULT 1

AA54375

ID AA54375 standard; Peptide; 44 AA.

AC AA54375;

DT 06-APR-2000 (first entry)

XX Amino acid sequence of a heliomyacin peptide.

XX Heliomyacin; polycysteine peptide; antifungal; antibacterial;

XX fungal infection; plant protection; animal protection;

XX transgenic plant; fungi; bacteria; Cercospora beticola;

XX Cladosporium herbarum; Fusarium culmorum; F. oxysporum;

XX Phytophthora carnationi.

XX Heliothis virescens.

XX Heliothis virescens.

XX Heliothis virescens.



(ENTO-) ENTOMED SA.  
 Dimarcq J, Legrain M, Menin L;  
 WPI; 2002-179779/23.  
 New peptide derivatives of heliomycin, useful as antibacterial and  
 antifungal agents in human or veterinary medicine and agriculture -  
 Claim 15; Page 57; 100pp; French.  
 The present invention relates to peptide derivatives of heliomycin with  
 antifungal and/or antibacterial activity. The present sequence is one  
 such peptide. The peptides are useful in human and veterinary medicine  
 and in plant protection, for control of bacterial and fungal infections.  
 An antifungal and/or antibacterial composition containing at least one  
 peptide is also claimed.  
 Sequence 44 AA;  
 Alignment Scores:  
 Pred. No.: 6.42e-26 Length: 44  
 Score: 267.00 Matches: 44  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 82.41% Indels: 0  
 Gaps: 23  
 DB: 0  
 JS-09-673-274B-2 (1-169) x ABB76868 (1-44)  
 1 CATAGCTTATCGGTTCTCGGTGGGGTCTGTGAACACTCCGATTGCAACGGT 60  
 1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20  
 61 GAGTGCAGAGAGGGGTTACAGGGTGGTCACTGCGGTTCCTTCGTTACGTTGAACCTGC 120  
 21 GluCysLysArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40  
 121 TGGTGGGAGACT 132  
 41 TrpCysGluThr 44  
 RESULT 4  
 AAY54374  
 D AAY54374 standard; Peptide: 49 AA.  
 AC AAY54374;  
 QY 06-APR-2000 (first entry)  
 DT Amino acid sequence of a fusion peptide of MF-alpha/heliomycin.  
 DE Mat alpha factor; MF-alpha; heliomycin; polycysteine peptide;  
 KW antifungal; antibacterial; fungal infection; plant protection;  
 QW animal protection; transgenic plant; fungi; bacteria;  
 CW Cercospora beticola; Cladosporium herbarum; Fusarium culmorum;  
 QW F. oxysporum; Phytophthora cinnamomi.  
 QX Chimeric - Saccharomyces cerevisiae.  
 DS Chimeric - Heliothis virescens.  
 QY Key Location/Qualifiers  
 DT Peptide 1.5  
 /note= "residues derived from the yeast Mat-alpha  
 (MF-alpha) factor"  
 DT Peptide 6.49  
 /note= "heliomycin residues"  
 FR2777568-A1.  
 QX 22-OCT-1999.  
 QY 15-APR-1998; 98FR-0004933.

XX 15-APR-1998; 98FR-0004933.  
 XX (RHON ) RHONE-POULENC AGROCHIMIE.  
 PA WPI; 2000-108532/10.  
 DR N-PSDB; AA245691.  
 XX New polycysteine peptides, designated heliomycin, with antifungal or  
 PT antibacterial activity, useful in medicine or plant protection.  
 XX Claim 13; Page 31; 46pp; French.  
 PS The present sequence represents a fusion peptide of 5 amino acids  
 CC derived from the Mat alpha factor (MF-alpha) and heliomycin. The  
 CC attachment of the MF-alpha sequence to the N-terminal aids secretion  
 CC of the heliomycin peptide. Heliomycin is a polycysteine peptide with  
 CC antifungal and antibacterial activity. Isolated from the lepidopteron  
 CC Heliothis virescens. Heliomycin peptides contain the sequence given in  
 CC AAY54378. The heliomycin peptides are used as pharmaceuticals for  
 CC treating or preventing fungal infections in humans and animals. They  
 CC may also be used as antifungal agents for plant protection. Nucleic  
 CC acid encoding the heliomycin peptides is used to generate transgenic  
 CC plants that are resistant to some fungi and bacteria (specifically  
 CC Cercospora beticola, Cladosporium herbarum, Fusarium culmorum,  
 CC F. oxysporum and Phytophthora cinnamomi), and for recombinant production  
 CC of the peptides.  
 XX Sequence 49 AA;  
 SQ Alignment Scores:  
 Pred. No.: 6.54e-26 Length: 49  
 Score: 267.00 Matches: 44  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 82.41% Indels: 0  
 Gaps: 21  
 DB: 0  
 US-09-673-274B-2 (1-169) x AAY54374 (1-49)  
 QY 1 GATAGCTTATCGGTTCTCGGTGGGGTCTGTGAACACTCCGATTGCAACGGT 40  
 DB 6 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 25  
 QY 61 GAGTGCAGAGAGGGGTTACAGGGTGGTCACTGCGGTTCCTTCGTTACGTTGAACCTGC 120  
 DB 26 GluCysLysArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 45  
 QY 121 TGGTGGGAGACT 132  
 DB 46 TrpCysGluThr 49  
 RESULT 5  
 AAY54376  
 ID AAY54376 standard; Peptide: 74 AA.  
 XX AAY54376;  
 AC AAY54376;  
 QX 06-APR-2000 (first entry)  
 DT A fusion peptide of PR-lalpha signal peptide/heliomycin.  
 DE PR-lalpha; heliomycin; polycysteine peptide; antifungal; antibacterial;  
 KW fungal infection; plant protection; animal protection; transgenic plant;  
 QW fungi; bacteria; Cercospora beticola; Cladosporium herbarum;  
 QW Fusarium culmorum; F. oxysporum; Phytophthora cinnamomi.  
 QX Chimeric - Nicotiana sp.  
 OS Chimeric - Heliothis virescens.  
 QX Key Location/Qualifiers  
 DT Peptide 1.30  
 /note= "tobacco PR-lalpha signal peptide"

FT Peptide 31..74  
 FT /note= "heliomycin"  
 XX  
 XX  
 PN FR277568-A1.  
 XX  
 PD 22-OCT-1999.  
 XX  
 XX 15-APR-1998; 98FR-0004933.  
 XX  
 XX 15-APR-1998; 98FR-0004933.  
 PR  
 XX (RHON ) RHONE-POULENC AGROCHIMIE.  
 PA  
 XX WPI: 2000-108532/10.  
 XX N-PSCH; AA245693.  
 DR  
 XX  
 XX New polycysteine peptides, designated heliomycin, with antifungal or  
 PT antibacterial activity, useful in medicine or plant protection .  
 PT  
 XX  
 XX Claim 13; Page 32; 46pp; French.  
 PS  
 XX  
 XX The present sequence represents a fusion peptide of the tobacco  
 CC PR-1alpha signal peptide and heliomycin. The attachment of the  
 CC PR-1alpha sequence to the N-terminal aids secretion of the heliomycin  
 CC peptide. Heliomycin is a polycysteine peptide with antifungal and  
 CC antibacterial activity, isolated from the lepidopteron *Heliothis*  
 CC virescens. Heliomycin peptides contain the sequence given in  
 CC AA554378. The heliomycin peptides are used as pharmaceuticals for  
 CC treating or preventing fungal infections in humans and animals. They  
 CC may also be used as antifungal agents for plant protection. Nucleic  
 CC acid encoding the heliomycin peptides is used to generate transgenic  
 CC plants that are resistant to some fungi and bacteria (specifically  
 CC *Cercospora beticola*, *Cladosporium herbarum*, *Fusarium culmorum*,  
 CC *F. oxysporum* and *Phytophthora cinnamomi*), and for recombinant production  
 CC of the peptides.  
 XX  
 XX Sequence 74 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 7,01e-26 Length: 74  
 Score: 267.00 Matches: 44  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 82.41% Indels: 0  
 DB: 21 Gaps: 0  
 US-09-673-274B-2 (1-169) x AAY54376 (1-74)  
 QY 1 GATAAGCTTATCGGTTCTCGGTGGGTGGTGGTGAACCTACACTCCGATTCGAACGGT 60  
 Db 31 AsplysLeuileGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 50  
 QY 61 GAGTGAAGAGGAGGGTTACAAGGGTGGTCACTGGGTTCTTCGCTACGTAAGTAAGTGC 120  
 Db 51 GlucysLysArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 70  
 QY 121 TGGTCCGAGACT 132  
 Db 71 TrpCysGluThr 74  
 RESULT 6  
 ABB76895  
 ID ABB76895 standard; Protein; 44 AA.  
 AC ABB76895;  
 XX  
 XX 03-JUL-2002 (first entry)  
 DT  
 XX Antifungal/antibacterial peptide #2.  
 DE  
 XX Antifungal; antibacterial; fungicide; heliomycin; human medicine;  
 KW veterinary medicine; plant protection; bacterial infection;  
 KW fungal infection.

XX Synthetic.  
 OS  
 XX WO200206324-A2.  
 PN  
 XX 24-JAN-2002.  
 PD  
 XX 05-JUL-2001; 2001WO-FR02164.  
 PF  
 XX 13-JUL-2000; 2000FR-0009248.  
 XX  
 PR 19-SEP-2000; 2000FR-0011949.  
 PR  
 XX (ENTO-) ENTOMED SA.  
 XX  
 XX Dimarcq J, Legrain M, Menin L;  
 XX WPI: 2002-179779/23.  
 DR  
 XX New peptide derivatives of heliomycin, useful as antibacterial and  
 PT antifungal agents in human or veterinary medicine and agriculture .  
 PT  
 XX  
 XX Disclosure; Page 86; 100pp; French.  
 PS  
 XX The present invention relates to peptide derivatives of heliomycin with  
 CC antifungal and/or antibacterial activity. The present sequence is one  
 CC such peptide. The peptides are useful in human and veterinary medicine  
 CC and in plant protection, for control of bacterial and fungal infections.  
 CC An antifungal and/or antibacterial composition containing at least one  
 CC peptide is also claimed.  
 XX  
 XX Sequence 44 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 8,65e-26 Length: 44  
 Score: 266.00 Matches: 43  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 97.73% Mismatches: 0  
 Query Match: 82.10% Indels: 0  
 DB: 23 Gaps: 0  
 US-09-673-274B-2 (1-169) x ABB76895 (1-44)  
 QY 1 GATAAGCTTATCGGTTCTCGGTGGGTGGTGGTGAACCTACACTTCGATTCGAACGGT 60  
 Db 1 AsplysLeuileGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20  
 QY 61 GAGTGAAGAGGAGGGTTACAAGGGTGGTCACTGGGTTCTTCGCTACGTAAGTAAGTGC 120  
 Db 21 GlucysLysArgGlyTyrLysGlyGlyHisCysGlySerPheAlaAsnValAsnCys 40  
 QY 121 TGGTCCGAGACT 132  
 Db 41 TrpCysGluThr 44  
 RESULT 7  
 ABB76863  
 ID ABB76863 standard; Protein; 44 AA.  
 XX  
 XX ABB76863;  
 XX  
 XX 03-JUL-2002 (first entry)  
 DT  
 XX Antifungal/antibacterial peptide pEM38.  
 DE  
 XX Antifungal; antibacterial; fungicide; heliomycin; human medicine;  
 KW veterinary medicine; plant protection; bacterial infection;  
 KW fungal infection.  
 XX  
 XX Synthetic.  
 OS  
 XX WO200206324-A2.  
 PN  
 XX 24-JAN-2002.  
 PD





PT New peptide derivatives of heliomicine, useful as antibacterial and  
 PT antifungal agents in human or veterinary medicine and agriculture  
 XX  
 XX Claim 15; Page 57; 100pp; French.  
 XX  
 XX The present invention relates to peptide derivatives of heliomicine with  
 CC antifungal and/or antibacterial activity. The present sequence is one  
 CC such peptide. The peptides are useful in human and veterinary medicine  
 CC and in plant protection, for control of bacterial and fungal infections.  
 CC An antifungal and/or antibacterial composition containing at least one  
 CC peptide is also claimed.  
 XX  
 XX Sequence 44 AA;  
 Alignment Scores:  
 Pred. No.: 1,57e-25 Length: 44  
 Score: 264.00 Matches: 43  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 97.73% Mismatches: 0  
 Query Match: 81.48% Indels: 0  
 DB: 23 Gaps: 0  
 US-09-673-274b-2 (1-169) x ABB76872 (1-44)  
 QY : GATAAGCTTATCGGTTCTCGGTGGGGTGTGTAACACTACACTCCGATTCCAAACGGT 60  
 Db :  
 1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20  
 61 GAGTGCAGAGAGGGTTACAGGGTGGTCACTGGCGGTCTCTCGCTAAGCTGAACCTGC 120  
 Db :  
 21 GluCysLysArgArgGlyTyrLysGlyHisCysGlySerPheAlaAsnValAsnCys 40  
 QY 121 TGGTGGCGAGACT 132  
 Db :  
 41 TrpCysGlnThr 44  
 RESULT 10  
 ABB76896  
 ID ABB76896 standard; Protein; 44 AA.  
 XX  
 AC ABB76896;  
 XX  
 DT 03-JUL-2002 (first entry)  
 XX  
 DE Antifungal/antibacterial peptide #3  
 XX  
 KW Antifungal; antibacterial; fungicide; heliomicine; human medicine;  
 KW veterinary medicine; plant protection; bacterial infection;  
 KW fungal infection.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200206324-A2.  
 XX  
 PD 24-JAN-2002.  
 XX  
 PF 05-JUL-2001; 2001WO-FR02164.  
 XX  
 PR 13-JUL-2000; 2000FR-0009248.  
 PR 19-SEP-2000; 2000FR-0011949.  
 XX  
 PA (ENTC-) ENTOMED SA.  
 XX  
 PI Dimarcq J, Legrain M, Menin L;  
 XX  
 DR WP1; 2002-179779/23.  
 XX  
 XX New peptide derivatives of heliomicine, useful as antibacterial and  
 PT antifungal agents in human or veterinary medicine and agriculture  
 XX  
 XX Disclosure; Page 88; 100pp; French.  
 XX  
 XX The present invention relates to peptide derivatives of heliomicine with

CC antifungal and/or antibacterial activity. The present sequence is one  
 CC such peptide. The peptides are useful in human and veterinary medicine  
 CC and in plant protection, for control of bacterial and fungal infections.  
 CC An antifungal and/or antibacterial composition containing at least one  
 CC peptide is also claimed.  
 XX  
 XX Sequence 44 AA;  
 Alignment Scores:  
 Pred. No.: 1,57e-25 Length: 44  
 Score: 264.00 Matches: 43  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 97.73% Mismatches: 0  
 Query Match: 81.48% Indels: 0  
 DB: 23 Gaps: 0  
 US-09-673-274b-2 (1-169) x ABB76896 (1-44)  
 QY 1 GATAAGCTTATCGGTTCTCGGTGGGGTGTGTAACACTACACTCCGATTCCAAACGGT 60  
 Db :  
 1 AspLysLeuIleGlySerCysValTrpGlyAlaValAsnTyrThrSerAspCysAsnGly 20  
 61 GAGTGCAGAGAGGGTTACAGGGTGGTCACTGGCGGTCTCTCGCTAAGCTGAACCTGC 120  
 Db :  
 21 GluCysLysArgArgGlyTyrLysGlyHisCysGlySerPheAlaAsnValAsnCys 40  
 QY 121 TGGTGGCGAGACT 132  
 Db :  
 41 TrpCysGlnThr 44  
 RESULT 11  
 ABB76862  
 ID ABB76862 standard; Protein; 44 AA.  
 XX  
 AC ABB76862;  
 XX  
 DT 03-JUL-2002 (first entry)  
 XX  
 DE Antifungal/antibacterial peptide pEM37.  
 XX  
 KW Antifungal; antibacterial; fungicide; heliomicine; human medicine;  
 KW veterinary medicine; plant protection; bacterial infection;  
 KW fungal infection.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200206324-A2.  
 XX  
 PD 24-JAN-2002.  
 XX  
 PF 05-JUL-2001; 2001WO-FR02164.  
 XX  
 PR 13-JUL-2000; 2000FR-0009248.  
 PR 19-SEP-2000; 2000FR-0011949.  
 XX  
 PA (ENTO-) ENTOMED SA.  
 XX  
 PI Dimarcq J, Legrain M, Menin L;  
 XX  
 DR WP1; 2002-179779/23.  
 XX  
 XX New peptide derivatives of heliomicine, useful as antibacterial and  
 PT antifungal agents in human or veterinary medicine and agriculture  
 XX  
 XX Claim 15; Page 57; 100pp; French.  
 XX  
 XX The present invention relates to peptide derivatives of heliomicine with  
 CC antifungal and/or antibacterial activity. The present sequence is one  
 CC such peptide. The peptides are useful in human and veterinary medicine  
 CC and in plant protection, for control of bacterial and fungal infections.  
 CC An antifungal and/or antibacterial composition containing at least one  
 CC peptide is also claimed.  
 XX

2 Sequence 44 AA;  
 Alignment Scores:  
 red. No.: 2.84e-25 Length: 44  
 core: 262.00 Matches: 43  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 97.73% Mismatches: 0  
 Query Match: 80.86% Indels: 0  
 B: 23 Gaps: 0

S-09-673-274B-2 (1-169) x ABB76862 (1-44)  
 Y 1 GATAAGCTTATCGGTTCTCGTGTGGGTGCTGTGAACACTACACTTCGCGATTGCAACGGT 60  
 b 1 AsnLysLeulleGlySerCysValTrpGlyAlaValAsnTrpThrSerAspCysAsnGly 20  
 Y 61 GAGTGAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCTTCGCTAACGTAACGTAACGTC 120  
 b 21 GluCysLysArgArgGlyTyrLysGlyGlyH: sCysGlySerPheAlaAsnValAsnCys 40  
 Y 121 TGGTCCGAGACT 132  
 b 41 TrpCysGluThr 44

RESULT 12  
 BB76865  
 D ABB76865 standard; Protein; 44 AA.  
 X  
 X ABB76865;  
 X  
 X 03-JUL-2002 (first entry)  
 X  
 X Antifungal/antibacterial peptide pEM42.  
 X  
 X Antifungal; antibacterial; fungicide; heliomycin; human medicine;  
 X veterinary medicine; plant protection; bacterial infection;  
 X fungal infection.  
 X Synthetic.  
 X  
 X WO200206324-A2.  
 X  
 X 24-JAN-2002.  
 X  
 X 05-JUL-2001; 2001WO-FR02164.  
 X  
 X 13-JUL-2000; 2000FR-0009248.  
 X  
 X 19-SEP-2000; 2000FR-0011949.  
 X  
 X (ENTO-) ENTOMED SA.  
 X  
 X Dimarcq J, Legrain M, Menin L;  
 X  
 X WPI; 2002-179779/23.  
 X  
 X New peptide derivatives of heliomycin, useful as antibacterial and  
 X antifungal agents in human or veterinary medicine and agriculture  
 X  
 X Claim 15; Page 57; 100pp; French.  
 X  
 X The present invention relates to peptide derivatives of heliomycin with  
 X antifungal and/or antibacterial activity. The present sequence is one  
 X such peptide. The peptides are useful in human and veterinary medicine  
 X and in plant protection, for control of bacterial and fungal infections.  
 X An antifungal and/or antibacterial composition containing at least one  
 X peptide is also claimed.  
 X  
 X Sequence 44 AA;  
 Alignment Scores:  
 red. No.: 2.84e-25 Length: 44  
 core: 262.00 Matches: 43  
 Percent Similarity: 97.73% Conservative: 0

Best Local Similarity: 97.73% Mismatches: 1  
 Query Match: 80.86% Indels: 0  
 DB: 23 Gaps: 0

US-09-673-274B-2 (1-169) x ABB76865 (1-44)  
 QY 1 GATAAGCTTATCGGTTCTCGTGTGGGTGCTGTGAACACTACACTTCGCGATTGCAACGGT 60  
 Db 1 AsnLysLeulleGlySerCysValTrpGlyAlaValAsnTrpThrSerAspCysAsnGly 20  
 QY 61 GAGTGAAGAGGAGGGGTTACAAGGGTGGTCACTGCGGTTCTTCGCTAACGTAACGTAACGTC 120  
 Db 21 GluCysLysArgArgGlyTyrLysGlyGlyH: sCysGlySerPheAlaAsnValAsnCys 40  
 QY 121 TGGTCCGAGACT 132  
 Db 41 TrpCysGluThr 44

RESULT 13  
 ABB76867  
 ID ABB76867 standard; Protein; 44 AA.  
 XX  
 AC ABB76867;  
 XX  
 DT 03-JUL-2002 (first entry)  
 XX  
 DE Antifungal/antibacterial peptide pEM22.  
 XX  
 KW Antifungal; antibacterial; fungicide; heliomycin; human medicine;  
 KW veterinary medicine; plant protection; bacterial infection;  
 KW fungal infection.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200206324-A2.  
 XX  
 PD 24-JAN-2002.  
 XX  
 PF 05-JUL-2001; 2001WO-FR02164.  
 XX  
 PR 13-JUL-2000; 2000FR-0009248.  
 PR 19-SEP-2000; 2000FR-0011949.  
 XX  
 PA (ENTO-) ENTOMED SA.  
 XX  
 PI Dimarcq J, Legrain M, Menin L;  
 XX  
 DR WPI; 2002-179779/23.  
 XX  
 PT New peptide derivatives of heliomycin, useful as antibacterial and  
 PT antifungal agents in human or veterinary medicine and agriculture  
 XX  
 PS Claim 15; Page 57; 100pp; French.  
 XX  
 CC The present invention relates to peptide derivatives of heliomycin with  
 CC antifungal and/or antibacterial activity. The present sequence is one  
 CC such peptide. The peptides are useful in human and veterinary medicine  
 CC and in plant protection, for control of bacterial and fungal infections.  
 CC An antifungal and/or antibacterial composition containing at least one  
 CC peptide is also claimed.  
 XX  
 SQ Sequence 44 AA;  
 Alignment Scores:  
 red. No.: 2.84e-25 Length: 44  
 core: 262.00 Matches: 43  
 Percent Similarity: 97.73% Conservative: 0  
 Best Local Similarity: 97.73% Mismatches: 1  
 Query Match: 80.86% Indels: 0  
 DB: 23 Gaps: 0

US-09-673-274B-2 (1-169) x ABB76867 (1-44)

Db	21	GlucCysLysArgGlyTyrLysGlyHisCysGlySerPheLeuAsnValAsnCys	40
QY	121	TGTCGCGAGACT 132	
Db	41	TrpCysGluThr 44	
RESULT 15			
ID	ABB76871		
DD	ABB76871	standard; Protein; 44 AA.	
XX	ABB76871		
AC	XX		
XX	XX		
DT	03-JUL-2002	(first entry)	
XX	XX		
DE	XX	Antifungal/antibacterial peptide PEM7.	
XX	XX		
KW	XX	Antifungal; antibacterial; fungicide; heliomicine; human medicine;	
KW	XX	veterinary medicine; plant protection; bacterial infection;	
KW	XX	fungal infection.	
OS	XX	Synthetic.	
XX	XX		
PN	WO200206324-A2.		
XX	XX		
PD	24-JAN-2002.		
XX	XX		
XX	35-JUL-2001; 2001WO-PRO2164.		
XX	XX		
PR	13-JUL-2000; 2000FR-0009248.		
PR	19-SEP-2000; 2000FR-001949.		
XX	XX		
PA	(ENTC-) ENTOMED SA.		
XX	XX		
PI	Dimarcq J, Legrain M, Wenin L;		
XX	XX		
DR	WPI; 2002-179779/23.		
XX	XX		
PT	XX	New peptide derivatives of heliomicine, useful as antibacterial and	
PT	XX	antifungal agents in human or veterinary medicine and agriculture	
PS	XX	Claim 15; Page 57; 100pp; French.	
XX	XX		
CC	XX	The present invention relates to peptide derivatives of heliomicine with	
CC	XX	antifungal and/or antibacterial activity. The present sequence is one	
CC	XX	such peptide. The peptides are useful in human and veterinary medicine	
CC	XX	and in plant protection, for control of bacterial and fungal infections.	
CC	XX	An antifungal and/or antibacterial composition containing at least one	
CC	XX	peptide is also claimed.	
SQ	XX	Sequence 44 AA;	
Alignment Scores:			
Pred. No.:	2.84e-25	Length:	44
Score:	262.00	Matches:	43
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.86%	Indels:	0
DB:	23.	Gaps:	0
US-09-673-274B-2 (1-169) x ABB76871 (1-44)			
QY	1	GATACCTTATCGGTTCTCGGTGGGGTCTGTGAACACTCTCCGATTGCAACGGT	60
Db	1	ApbLysLeuileGlySerCyavalTrpGlyalalalanTyThrSeraspCysAsnGly	20
QY	61	GAGTGCAGAGAGGGGTTACAGGGTGGTGCATCGCGGTTCCCTTCCGTAACGTGAAC	120
Db	21	GlucCysLysArgGlyTyrLysGlyHisCysGlySerPheAlaAsnValAsnCys	40
QY	121	TGTCGCGAG 129	
Db	41	TrpCysGlu 43	

Search completed: October 16, 2003, 17:10:35  
Job time : 52.2159 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

MM nucleic - protein search, using frame_plus_n2p model

un on:      October 16, 2003, 17:12:07 ; Search time 35.0914 Seconds
           (without alignments)
           1578.850 Million cell updates/sec

title:
effect score: 324
sequence: 1 gataagcttattcgttcctcg.....gaacgtgtcgacggatccgg 169

coring table:
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

searched: 609560 seqs, 163197102 residues

Total number of hits satisfying chosen parameters: 1219120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
MODEL=frame_n2p.model -DEV=xlp
Q=/cgn2_1/USPro.spool/US09673274/runat_16102003_170630_2743/app.query.fasta_1.654
DB=Published Applications AA -OFMT=tastan -SUFFIX=rapb -MINMATCH=0.1
LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
TRANS=human40.cdi -LIST=45 -DOCAUGN=200 -THR SCORE=pct -THR MAX=100
THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0
MAXLEN=2000000000 -USER=US09673274@cgn2_1.06 -RUNAT=16102003_170630_2743
NCPU=6 -ICPU=3 -NO MAP -LARGECUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=C.5
FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result  #      Score      Query      Match      Length DB ID      Description
-----

```

```
1 232 71.6 44 15 US-10-125-258-119
2 232 71.6 73 15 US-10-125-258-117
3 79.5 24.5 65 15 US-10-264-480-13
4 79.5 24.5 66 15 US-10-264-480-11
5 76.5 23.6 61 15 US-10-264-480-9
6 75 23.1 61 9 US-09-623-481-2
7 74.5 23.0 44 12 US-10-180-247-4
8 74.5 23.0 69 12 US-10-180-247-6
9 74.5 23.0 70 12 US-10-180-247-2
10 71.5 22.1 61 15 US-10-264-480-5
11 71.5 22.1 61 15 US-10-264-480-7
12 71.5 22.1 61 15 US-10-264-480-8
13 71 21.9 47 15 US-10-178-213-411
14 71 21.9 77 15 US-10-178-213-410
15 70.5 21.8 47 15 US-10-178-213-387
16 70.5 21.8 65 15 US-10-264-480-14
17 70.5 21.8 84 15 US-10-178-213-386
18 70.5 21.8 87 15 US-10-264-480-10
19 69.5 21.5 84 15 US-10-264-480-18
20 69 21.3 47 15 US-10-178-213-408
21 69 21.3 47 15 US-10-178-213-444
22 69 21.3 77 15 US-10-178-213-407
23 69 21.3 78 15 US-10-178-213-443
24 68.5 22.6 173 15 US-10-156-761-7787
25 68 21.0 47 15 US-10-178-213-396
26 68 21.0 76 15 US-10-178-213-395
27 67.5 20.8 87 15 US-10-264-480-12
28 67.5 22.3 433 15 US-10-08-872-208
29 67 20.7 45 15 US-10-178-213-468
30 67 20.7 72 15 US-10-178-213-567
31 67 20.7 86 15 US-10-264-480-22
32 67 20.7 86 15 US-10-264-480-23
33 66.5 20.5 47 15 US-10-178-213-309
34 66.5 20.5 47 15 US-10-178-213-339
35 66.5 20.5 58 15 US-10-264-480-4
36 66.5 20.5 62 14 US-10-044-359-21
37 66.5 20.5 66 15 US-10-264-480-15
38 66.5 20.5 77 15 US-10-178-213-308
39 66.5 20.5 84 15 US-10-178-213-338
40 66.5 21.9 810 15 US-10-156-761-10081
41 66 21.8 198 10 US-09-764-868-677
42 66 21.8 307 10 US-09-764-868-1700
43 66 21.8 312 9 US-09-925-302-783
44 65.5 20.2 44 15 US-10-178-213-147
45 65.5 20.2 58 15 US-10-264-480-3
```

## ALIGNMENTS

```
RESULT 1
US-10-125-258-119
; Sequence 119, Application US/10125258
; Publication No. US2001028926A1
; GENERAL INFORMATION:
; APPLICANT: Altier, Daniel J.
; APPLICANT: Hermann, Rafael
; APPLICANT: Lu, Albert J.
; APPLICANT: McCutchen, Billy F.
; APPLICANT: Presnail, James K.
; APPLICANT: Weaver, Jarine L.
; APPLICANT: Wong, James F. H.
; TITLE OF INVENTION: Antimicrobial Polypeptides and Their
; USES
; FILE REFERENCE: 35718/246215
; CURRENT APPLICATION NUMBER: US/10/125,258
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 60/285,355
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119
; LENGTH: 44
; TYPE: PRT
```

Alignment Scores:	
Pred. No.:	0.162
Score:	79.50
Percent Similarity:	53.85%
Best Local Similarity:	43.59%
Query Match:	24.54%
Indels:	7
Mismatches:	11
Conservative:	4
Matches:	17
Length:	66





```

; GENERAL INFORMATION:
; APPLICANT: Derose, Richard
; APPLICANT: Freysinnet, Georges
; APPLICANT: Hoffman, Jules
; TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
; TITLE OF INVENTION: Vector Containing It and Production of Disease-Resistant
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: A32889-PCT-USA-A 072667.0152
; CURRENT APPLICATION NUMBER: US/10/180,247
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 09/480,251
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: PCT/FR98/01462
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: FR97/09,115
; PRIOR FILING DATE: 1997-07-11
; PRIOR APPLICATION NUMBER: FR97/09,663
; PRIOR FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ubiquitin drosomycin fusion peptide
US 10 180-247-6

Alignment Scores:
Pred. No.: 0.699 Length: 69
Score: 74.50 Matches: 25
Percent Similarity: 42.50% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 16
Query Match: 22.99% Indels: 7
DB: 12 Gaps: 3

US-09-673-274B-2 (1-169) x US-10-180-247-6 (1-69)
QY 13 GGTCTCTGC---GTGTGGGTGCTGTGAACACTCCGATTGCAACGGTGAGTGCAG 69
Db 34 GlyProCysAlaValTrpAsp -----AsnGluThrCysArgValCysLys 49
QY 70 AGAGGGGTACAGGGTGTCTGCTGCGGTTCTCGCTTAACGTGAACCTGCTGGTGGAG 129
Db 50 GluGluGlyArgSerSerGlyHisCysSer-----ProSerLeuLysCysTrpCysGlu 67

RESULT 9
US-10-180-247-2
; Sequence 2, Application US/10180247
; Publication No. US20030167519A.
; GENERAL INFORMATION:
; APPLICANT: Derose, Richard
; APPLICANT: Freysinnet, Georges
; APPLICANT: Hoffman, Jules
; TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
; TITLE OF INVENTION: Vector Containing It and Production of Disease-Resistant
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: A32889-PCT-USA-A 072667.0182
; CURRENT APPLICATION NUMBER: US/10/180,247
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 09/480,251
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/FR98/01462
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: FR97/09,115
; PRIOR FILING DATE: 1997-07-11
; PRIOR APPLICATION NUMBER: FR97/09,663
; PRIOR FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 70
; TYPE: PRT

```

```

; ORGANISM: Drosophila melanogaster
US-10-180-247-2

Alignment Scores:
Pred. No.: 0.7 Length: 70
Score: 74.50 Matches: 15
Percent Similarity: 42.50% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 16
Query Match: 22.99% Indels: 7
DB: 12 Gaps: 3

US-09-673-274B-2 (1-169) x US-10-180-247-2 (1-70)
QY 13 GGTCTCTGC---GTGTGGGTGCTGTGAACACTCCGATTGCAACGGTGAGTGCAG 69
Db 35 GlyProCysAlaValTrpAsp -----AsnGluThrCysArgValCysLys 50
QY 70 AGAGGGGTACAGGGTGTCTGCTGCGGTTCTCGCTTAACGTGAACCTGCTGGTGGAG 129
Db 51 GluGluGlyArgSerSerGlyHisCysSer-----ProSerLeuLysCysTrpCysGlu 68

RESULT 10
US-10-264-480-6
; Sequence 6, Application US/10264480
; Publication No. US2003013892A1.
; GENERAL INFORMATION:
; APPLICANT: Hammock, Bruce D.
; APPLICANT: Inceoglu, Bora
; TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
; FILE REFERENCE: JCAL256
; CURRENT APPLICATION NUMBER: US/10/264,480
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/393,070
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/327,602
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Buthacus arenicola
US-10-264-480-6

Alignment Scores:
Pred. No.: 1.65 Length: 61
Score: 71.50 Matches: 14
Percent Similarity: 55.26% Conservative: 7
Best Local Similarity: 36.84% Mismatches: 12
Query Match: 22.07% Indels: 5
DB: 15 Gaps: 2

US-09-673-274B-2 (1-169) x US-10-264-480-6 (1-61)
QY 16 TCCTGCTGTGGGTGCTGTGAACACTCCGATTGCAACGGTGAGTGCAGAGAGG 75
Db 13 SerCysLeuPheGly-----AsnGluGlyCysAspLysGluCysLysAlaTyr 28
QY 76 GGTTCACAGGGTGTCTGCTGCGGTTCTCGCTTAACGTGAACCTGCTGGTGGAG 129
Db 29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 11
US-10-264-480-7
; Sequence 7, Application US/10264480
; Publication No. US2003013892A1.
; GENERAL INFORMATION:
; APPLICANT: Hammock, Bruce D.
; APPLICANT: Inceoglu, Bora
; TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
; FILE REFERENCE: JCAL256

```







```
QY 100 AACCGCAGTGACACCCCTGTG-----AACCCCTCTCTTCTG-----ACT 62
DB 200 AAGAGAGA-----AlaCysArgGlyGlyLysProAlaGAspCysGlyCysArgTrpCys 217
QY 61 CACCGTTGCAATCGAAGGTGAGTTTCACAGCAGCAGCAGCAGG 17
DB 218 HIsArgCysArgArgProGlySerThrAlaGlyProAaAAsArg 232

RESULT 2
US-08-435-040-2
Sequence 2, Application US/08435040
Patent No. 5756340
GENERAL INFORMATION:
APPLICANT: Hammock, Bruce D.
APPLICANT: Herrmann, Rafael
APPLICANT: Moskowitz, Haim
TITLE OF INVENTION: Insect Control With Multiple Toxins
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Majestic, Parsons, Siebert & Hsue
STREET: Four Embarcadero Center, Suite 1450
CITY: San Francisco
STATE: California
COUNTRY: U.S.A.
ZIP: 94111-4121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DCS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,040
FILING DATE: 08-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Siebert, J. Suzanne
REGISTRATION NUMBER: 28,758
REFERENCE/DOCKET NUMBER: 2500.078US0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 362-5556
TELEFAX: (415) 362-5418
TELEX: 278638 MGPS
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-435-040-2

Alignment Scores:
Pred. No.: 0.0254 Length: 65
Score: 80.50 Matches: 17
Percent Similarity: 57.14% Conservative: 7
Best Local Similarity: 40.48% Mismatches: 9
Query Match: 24.85% Indels: 9
DB: 1 Gaps: 4

US-09-673-274B-2 (1-169) x US-08-435-040-2 (1-65)
QY 16 TCCTGCGTGTGG-----GGTGGTGTAACACTACACTTCCGATTGCAACGGTGAGTCAAG 69
DB 12 AsnCysValThrCysGlyAlaAsnSerTyr-----CysAsnThrGluCysThr 28
QY 70 AGGAGGGGTTACAAGGGTGTGCTACTGC-----GGTTCCTTCGCTAACGTGAAGTGC 120
DB 29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpPheGlyLysTyrGlyAsnAla----Cys 47
QY 121 TGGTGC 126
DB 48 TrpCys 49

RESULT 4
```

```
Db 48 TrpCys 49
RESULT 3
US-09-020-216-2
Sequence 2, Application: US/99020216
Patent No. 6162430
GENERAL INFORMATION:
APPLICANT: Hammock, Bruce D.
APPLICANT: Herrmann, Rafael
APPLICANT: Moskowitz, Haim
TITLE OF INVENTION: Insect Control With Multiple Toxins
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Majestic, Parsons, Siebert & Hsue
STREET: Four Embarcadero Center, Suite 1450
CITY: San Francisco
STATE: California
COUNTRY: U.S.A.
ZIP: 94111-4121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DCS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,216
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/435,040
FILING DATE: 08-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Siebert, J. Suzanne
REGISTRATION NUMBER: 28,758
REFERENCE/DOCKET NUMBER: 2500.078US0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 362-5556
TELEFAX: (415) 362-5418
TELEX: 278638 MGPS
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-09-020-216-2

Alignment Scores:
Pred. No.: 0.0254 Length: 65
Score: 80.50 Matches: 17
Percent Similarity: 57.14% Conservative: 7
Best Local Similarity: 40.48% Mismatches: 9
Query Match: 24.85% Indels: 9
DB: 3 Gaps: 4

US-09-673-274B-2 (1-169) x US-09-020-216-2 (1-65)
QY 16 TCCTGCGTGTGG-----GGTGGTGTAACACTACACTTCCGATTGCAACGGTGAGTCAAG 69
DB 12 AsnCysValThrCysGlyAlaAsnSerTyr-----CysAsnThrGluCysThr 28
QY 70 AGGAGGGGTTACAAGGGTGTGCTACTGC-----GGTTCCTTCGCTAACGTGAAGTGC 120
DB 29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpPheGlyLysTyrGlyAsnAla----Cys 47
QY 121 TGGTGC 126
DB 48 TrpCys 49

RESULT 4
```

FILING DATE: 17-JULY-1996  
APPLICATION NUMBER: US/08/256,933  
FILING DATE: 27-JULY-1994  
APPLICATION NUMBER: WO 93/15:08  
FILING DATE: 29-JAN-1993  
APPLICATION NUMBER: AJ EL0722  
FILING DATE: 31-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Shaw, Melissa A.  
REGISTRATION NUMBER: 38,301  
REFERENCE/DOCKET NUMBER: PFD 5C99/D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 510-231-1112  
TELEFAX: 510-231-1112  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 61 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Leivirus quinquestratus quinquestriatus

Alignment Scores:  
Pred. No.: 0.C812 Length: 61  
Score: 76.50 Matches: 15  
Percent Similarity: 55.26% Conservative: 6  
Best Local Similarity: 39.47% Mismatches: 12  
Query Match: 23.61% Indels: 5  
DB: Gaps: 2

US-09-673-274B-2 (1-169) x US-08-682-485A-25 (1-61)  
QY 16 TCTCCGTGGGTCGTGAATACACTTCCAGTTCCAAAGCGTGAGTCAAGAGG 129  
Db 13 SerCysLeuPheGly-----AsnGluGlyCysAsnLysGluCysLysSerTy 45  
QY 76 GTTACAAGGTCATCGGTCCTTCTCGCTAACGTGAACCTGCTGTGCAG 129  
Db 29 GlyGlySerTyrglytyrCysArprThrTrf---GlyLeuAlaCysfrtpCyseglu 45

RESULT 6  
US-08-451-472-4  
Sequence 4, Application US/C8451472  
Patent No. 5770192  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: BIOLOGICAL CONTROL AGENTS  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye PC  
STREET: 8th Floor, 1100 No. 5770192th Glebe Road  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/451.472  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION NUMBER: US 08/117,125  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mary J Wilson

-09-252-991A-32270  
Sequence 32270, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 32270  
LENGTH: 149  
TYPE: PRt  
ORGANISM: Pseudomonas aeruginosa

-09-252-991A-32270  
ignment Scores:  
ed. No.: 0.0607 Length: 149  
ore: 78.00 Matches: 18  
cent Similarity: 49.06% Conservative: 8  
st Local Similarity: 33.96% Mismatch: 21  
ery Match: 25.74% Indels: 6  
Gaps: 4

-09-673-274B-2 (1-169) x US-09-252-991A-32270 (1-149)  
169 CCGATCGTCGACGAGTCGCCGCGAGCTCTCAAGTCTGCCACCAGCATTCACGT 110  
14 ProSerProAlaSerArgArgArgArgCysSerProSerArgSerProThrGly--- 32  
109 TAGCGAAGCACCCCAGTGACCCCTTGTAACCCCTCTCTTGCACCTACCCGTTGCAAT 50  
33 ---ArgtyrThrArgAlaProTrpArgCys-----CysArgArgArgCysArg 47  
49 CGGAAGTGTCAGTCACAGCACCACCGACGAGCAACCGA 11  
48 ArgArgCysArgProArgArgProThrSerLysalaarg 60

USULT 5  
-08-682-485A-25  
Sequence 25, Application US/08682485A  
Patent No. 5763568  
GENERAL INFORMATION:  
APPLICANT: ATKINSON, RONALD K  
APPlicant: HOWDEN, MERLIN E.H.  
APPlicant: TYLER, MARGARET I.  
APPlicant: VONARX, EDWARD J  
TITLE OF INVENTION: Insecticidal Toxins Derived From  
TITLE OF INVENTION: Funnel Web (Atrox or Hadronyche Spiders)  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESS: Zeneca, Inc.  
STREET: 1200 South 47th Street  
CITY: Richmond  
STATE: California  
COUNTRY: USA  
ZIP: 94804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/682.485A  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION NUMBER: US/08/682.485

FILING DATE: 17-JULY-1996  
APPLICATION NUMBER: US/08/256,933  
FILING DATE: 27-JULY-1994  
APPLICATION NUMBER: WO 93/15:08  
FILING DATE: 29-JAN-1993  
APPLICATION NUMBER: AJ EL0722  
FILING DATE: 31-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Shaw, Melissa A.  
REGISTRATION NUMBER: 38,301  
REFERENCE/DOCKET NUMBER: PFD 5099/D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 510-231-1122  
TELEFAX: 510-231-1122  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 61 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Leivirus quinquestratus quinquestriatus

US-08-682-485A-25

Alignment Scores:

Pred. No.:	Score:	Length:
0.0812	76.50	61
Percent Similarity:	55.26%	Matches:
Best Local Similarity:	39.47%	Conservative:
Query Match:	23.61%	Mismatches:
DB:	1	Indels:
		Gaps:

US-09-673-274B-2 (1-169) x US-08-682-485A-25 (1-61)

QY 16 TCTCCGTGGGTCGTGAAGTACACTCCGATTCGACCGTGAGTGCAGAGGAGG 129  
||| ||||||||| ||||||| ||||||| |||||||  
Db 13 SerCysLeuPheGly-----AsnGluGlyCysAsnLysGluCysLysSerTy 45  
QY 76 GTTACAAGGTTGTCACCTGGCGTTCCTTCGCTAACGTGAACCTGCTGGTGGAG 129  
||| ||||||| ||||||| ||||||| |||||||  
Db 29 G.YglySerTyrglyTyrcysArpThrTrf---GlyLeuAlaCysfrpCysglu 45

RESULT 6  
US-08-451-472-4  
Sequence 4, Application US/C8451472  
Patent No. 5770192  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: BIOLOGICAL CONTROL AGENTS  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye PC  
STREET: 8th Floor, 1100 No. 5770192th Glebe Road  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/451.472  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION NUMBER: US 08/117,125  
FILING DATE: 24-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mary J Wilson

-09-252-991A-32270  
Sequence 32270, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 32270  
LENGTH: 149  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa

-09-252-991A-32270

ignment Scores:

Pred. No.:	Score:	Length:
0.0607	78.00	149
Percent Similarity:	49.06%	Matches:
Best Local Similarity:	33.96%	Conservative:
Query Match:	25.74%	Mismatches:
DB:	4	Indels:
		Gaps:

-09-673-274B-2 (1-169) x US-09-252-991A-32270 (1-149)

169 CCGATCGTCGACGACGTCGCCGCGAGCTCAAGTTCGCCACCAGCAGTTCACGT 110  
||| ||||||||| ||||||| ||||||| |||||||  
14 ProSerProAlaSerArgArgArgArgCysSerProSerArgSerProThrGly--- 32  
109 TAGCGAAGGAAACCCAGTGACCCCTTGTAACCCCTTCCTTCGCACTCACCGTTCGAAT 50  
||| ||||||| ||||||| ||||||| |||||||  
33 ---ArgTyThrArgAlaProTrpArgCys-----CysArgArgArgCysArg 47  
49 CGGAAGTGTAGTTCACAGCACCCCACGACGAGGAAACCGA 11  
||| ||||||| ||||||| ||||||| |||||||  
48 ArgArgCysArgProArgArgProThrSerLysalaArg 60

USULT 5  
-08-682-485A-25  
Sequence 25, Application US/08682485A  
Patent No. 5763568  
GENERAL INFORMATION:  
APPLICANT: ATKINSON, RONALD K  
APPLICANT: HOWDEN, MERLIN E.H.  
APPLICANT: TYLER, MARGARET I.  
APPLICANT: VONARX, EDWARD J  
TITLE OF INVENTION: Insecticidal Toxins Derived From  
TITLE OF INVENTION: Funnel Web (Attax or Hadronyche Spiders)  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Zeneca, Inc.  
STREET: 1200 South 47th Street  
CITY: Richmond  
STATE: California  
COUNTRY: USA  
ZIP: 94804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/682.485A  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION NUMBER: US/08/682.485









PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 26857

LENGTH: 397  
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa  
S-09-252-991A-26857

Alignment Scores:  
red. No.: 0.749 Length: 397  
core: 70.00 Matches: 20  
Percent Similarity: 51.02% Conservative: 5  
Best Local Similarity: 40.82% Mismatches: 23  
Very Match: 23.10% Indels: 1  
B: 4 Gaps: 0

S-09-673-274B-2 (1-169) x US-09-252-991A-26857 (1-397)

```
y 169 CCGATCCGTCGACACGTTCCGCTCGCGAGCTCTCAAGTCTCGACACGACGACGTTCCAGCT 110
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
b 270 ProSerArgSerThrArgLysProSerArgLysProSerArgLysProSerArgLysProSerArgLys 289
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
y 109 TAGCGAAGGAACCGACGTCGACCCCTTGTAAACCCCTCTCTTGCACCTCACCCTTGCAT 50
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
b 290 ArgArgArgSerSerSerSerGlyArgSerThrProSerTrpCysThrArg-ValAlaIle 309
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
y 49 CGAAGTGTAGTTCACAGACCCCA 25
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
b 309 eSerSerIleSerSerThrAlaPro 317
```

Search completed: October 16, 2003, 17:13:05  
Job time : 18.8439 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

DM nucleic - protein search, using frame\_plus\_n2p model

Run on: October 16, 2003, 17:06:37 ; Search time 16.2259 Seconds  
(without alignments)  
1564.690 Million cell updates/sec

Title: US-09-673-274B-2\_COPY\_1\_132

Perfect score: 267

Sequence: 1 gataagcttcctgctctg.....tgaactcctggtcgagact 132

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODEL=frame+ n2p.model -DEV=xlp  
O=/cgm2\_1/USPTO\_epool/US09673274/runat\_16102003\_170628\_2555/app\_query.fasta\_1.654  
DB=PIR\_76 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
OUTFWT=ptp -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
USER=US09673274 -SCGN 1 1 77 @runat\_16102003\_170628\_2555 -NCPU=6 -ICPU=3  
NO\_MMWP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR\_76:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	32.2	66	2	A29386 neurotoxin I-I - b
2	80.5	30.1	65	2	A59222 neurotoxin X - sco
3	79.5	29.8	66	2	A58869 crustacean-specific
4	79	29.6	85	2	alpha-toxin Bot XI
5	78.5	29.4	84	1	NTSR3A neurotoxin III pre
6	77.5	29.0	65	1	NTSR4L neurotoxin IV - Eg
7	76.5	28.7	61	2	B34123 depressant insect
8	76	28.5	66	1	NTSR4E neurotoxin M14 - l
9	75.5	28.3	65	1	NTSR1C neurotoxin I - bar
10	74.5	27.9	70	2	A58824 drosomycin precurs
11	74	27.7	59	2	C23727 neurotoxin V-5 - b
12	74	27.7	64	2	JC1321 neurotoxin IV - Sa
13	73	27.3	64	1	NTSR5L neurotoxin V - Egy
14	70.5	26.4	65	1	NTSR3C neurotoxin 3 - bar

15	70.5	26.4	87	2	JN0670 Na+-channel-blocki
16	69.5	26.0	84	1	NTSR2N Na+-channel-blocki
17	68.5	25.7	62	2	D59352 depressant insect-
18	68.5	25.7	66	1	NTSR9E neurotoxin N9 - le
19	68.5	25.7	83	1	NTSR1A neurotoxin I precu
20	68	25.5	64	1	NTSR5M proteinase inhibit
21	68	25.5	77	2	T14395 neurotoxin V - sco
22	67.5	25.3	61	2	B59352 depressant insect-
23	67.5	25.3	61	2	D59352 depressant insect-
24	67.5	25.3	83	1	NTSR12 neurotoxin I' proc
25	67.5	25.3	87	2	JN0669 Na+-channel-blocki
26	67	25.1	86	2	JN0671 Na+-channel-blocki
27	66.5	24.9	65	1	NTSR1B neurotoxin I - sco
28	66.5	24.9	66	1	NTSR2C neurotoxin 2 - bar
29	66	24.7	3672	2	T23433 hypothetical prote
30	66	24.7	3704	2	T17116 probable laminin a
31	65.5	24.5	61	2	A59006 excitatory insect
32	64.5	24.2	66	2	B23727 neurotoxin V-4 - b
33	64.5	24.2	85	2	A61616 depressant insect
34	64	24.0	77	2	S10578 protease inhibitor
35	64	24.0	78	2	S52634 gamma-thionin Ph
36	64	24.0	92	2	S43228 defensin - fruit f
37	63.5	23.8	85	2	A40472 depressant insect
38	63	23.6	64	2	A35940 neurotoxin alpha-I
39	63	23.6	85	2	A39306 alpha insect toxin
40	63	26.8	4753	1	A47437 LDL-receptor-relat
41	62	23.2	3712	2	S18253 laminin alpha-1 ch
42	61.5	23.0	65	1	NTSR2R neurotoxin II - sc
43	61.5	23.0	65	2	A38394 neurotoxin Aah I74
44	61.5	23.0	66	2	A27406 toxin Cssi - Mexi
45	61.5	23.0	84	2	JE0143 anti-mammals neuro

#### ALIGNMENTS

##### RESULT 1

A29386 neurotoxin III - scorpion (Buthus occitanus)

C/Species: Buthus occitanus mardochaei

C/Date: 21-May-1988 #sequence\_revision: 21-May-1988 #text\_change: 07-Feb-1997

C/Accession: A29386

R/Vargas, O.; Martin, M.F.; Rochat, H.

Eur. J. Biochem. 162, 589-599, 1987

A/Title: Characterization of six toxins from the venom of the Moroccan scorpion Buthus

A/Reference number: A29386; XUID:87161929; PMID:3104036

A/Accession: A29386

A/Molecule type: protein

A/Residues: 1-66 <VAR>

C/Superfamily: scorpion neurotoxin

C/Keywords: neurotoxin

##### Alignment Scores:

Pred. No.:	0.0247	Length:	66
Score:	86.00	Matches:	15
Percent Similarity:	55.00%	Conservative:	7
Best Local Similarity:	37.50%	Mismatches:	16
Query Match:	32.21%	Indels:	2
DB:	2	Gaps:	1

US-09 673-274B-2\_COPY\_1\_132 (1-132) x A29386 (1-66)

Qy 16 TCCTGGGTGGGGTGGTGTGAACCTACACTTCGGATTGCAACGGTGAGTGCAGAGGAGG 75

Db 11 AsnCysValThrHisCysPheProGlySerSerGlyCysAspThrLeuCysLysGlyLys 10

Qy 76 GGTACAGCGGTGGTCACTCGGTTCTTC-----GCTAACGCTGAACCTGGTGGAG 129

Db 31 GlyAlaThrSerGlyHisCysGlyPheLeuProGlySerGlyValAlaCysTrpCysAsp 50

##### RESULT 2

A59222

neurotoxin X - scorpion (Leiurus quinquestriatus)

C/Species: Leiurus quinquestriatus hebraeus

```
C>Date: 31 Dec-2001 #sequence_revision 31-Dec-2001 #text_change 19-Apr-2002
C:Accession: A59222
R:Corzo, G.; Escobas, P.; Nakajima, T.
submitted to the Protein Sequence Database, April 2000
A:Description: A procedure for the rapid purification of non commercial and new scorpion
A:Reference number: A59222
A:Accession: A59222
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-65 <COR>
C:Superfamily: scorpion neurotoxin
C:Keywords: amidated carboxyl end; neurotoxic; venom
F:13-64,17,37,23-47,27-49/Dissulfide bonds: #status predicted
F:65/Modified site: amidated carboxy end (Arg) #status predicted

Alignment Scores:
Pred. No.:      Length:      65
Score:          80.50
Percent Similarity: 57.14%
Best Local Similarity: 40.48%
Query Match:    30.15%
DB:             2           4

US_09 673-274B-2_COPY_1_132 (1-132) x A59222 (1-65)

QY       16 TCCTGGGTGG-----GTCGTGCAACTACACTCCGATTCGAACGGTGAGTGCNAG 69
DB       12 ASCcysvallyThrCysG-YAlaAsnSerTy.....CysAnthrG-uCysThr 28
QY       70 AGAGGGGGTTACAAGGTGCTCATCTOC-----GGTCCTTCGCTTAACGTGAAGACTGC 120
DB       29 LysAsnGlyAlaGluserG-yTYrCySGlnTrpPheGlyLysTYrGlyASrAla---Cys 47
QY       121 TCSTGCC 126
DB       45 TrpCys 49

RESULT :
A55869
Crustacean specific toxin 1 - scorpion (Centruroides limpidus)
C:Species: Centruroides limpidus limpidus
C>Date: 25-Jan-1996 #sequence_revision 05-Jun-1996 #text_change 17-Mar-1999
C:Accession: A55869
Biochemistry 33, 11135-11149, 1994
A:Author(s): Leclercq, F.; Ramirez, M.; Ramirez, A.N.P.; Balderas, C.; Possani, L.D.
A>Title: Primary and NMR three-dimensional structure determination of a novel crustacean
A:Reference number: A55869; PMID:95244424; PMID:7727365
A:Accession: A55869
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-66 <LEB>
C:Superfamily: scorpion neurotoxin
C:Keywords: disulfide bond; toxin
F:12-65,16-41,25-46,29-48/Dissulfide bonds: #status experimental

Alignment Scores:
Pred. No.:      Length:      66
Score:          79.50
Percent Similarity: 53.85%
Best Local Similarity: 43.59%
Query Match:    29.78%
DB:             2           3

US_09 673-274B-2_COPY_1_132 (1-132) x A55869 (1-66)

QY       19 TGCGTGTGGGTGCTGTGAACTACACTCCGATTGCAACGGTGAGTGCNAGAGAGGGGT 78
DB       16 CysPheIrpLeuGlyLYsAsn-----GLuSncCysAspLySGluCYsLYsA-LysAsn 33
QY       79 TACAAGGGT-----GGTCACGTGGGTTCCTTCGCTAACGTGAAGTGCCTGGTGCAG 129
DB       34 GInGlyGlySerTYrGlyTYrCysTySerPheAla-----CysTrpCysGlu 49
```

Biochemistry 30, 4814-4821, 1991

A;Title: Functional duality and structural uniqueness of depressant insect selective toxin A;Reference number: A38526; MUID:91230120; PMID:2029523

A;Accession: B38526

A;Molecule type: protein

A;Residues: 1-61 <ZL02>

A;Cross-references: CAS:130300-67 3

C;Superfamily: scorpion neurotoxin

C;Keywords: neurotoxin; venom

F;10-60,14-35,21-42,25-44/Disulfide bonds: #status predicted

Alignment Scores:

Pred. No.:	0.368	length:	61
Score:	76.50	Matches:	15
Percent Similarity:	55.26%	Conservative:	6
Best Local Similarity:	39.47%	Mismatches:	12
Query Match:	28.65%	Indels:	5
DB:	2	Gaps:	2

US-09-673-274B-2\_COPY\_1\_132 (1-132) x B34:23 (1-61)

Qy 16 TCCTGCGTGTGGGTCTGTGAACACTACACTCCGATTCGACGGTGAAGAGGAGG 75

Db 13 SerCysLeuPheGly.....AsnGluG:CysAsnLysGluCysLysSerTyr 28

Qy 76 GTTACAAAGGGTGTCACTGCGGTTCCTTCGCTACGTAAGTACTGCTGTGGCGAG 129

Db 29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 8

NTSR4E

neurotoxin M14 - lesser Asian scorpion

C;Species: Mesobuthus eupeus (lesser Asian scorpion)

C;Date: 30-Sep-1988 #sequence\_revision 30-Sep-1988 #text\_change 31-Dec-1993

C;Accession: J0020

R;Volkova, T.M.; Garsia, A.F.; Telezhinskaya, I.N.; Potapenko, N.A.; Grishin, E.V.

R;Boorg, Khim. 11, 1445-1456, 1985

A;Title: Study of neurotoxins from the venom of central Asian scorpion Buthus eupeus

A;Reference number: A94652; MUID:86130759; PMID:4091860

A;Accession: J0020

A;Molecule type: protein

A;Residues: 1-66 <VOL>

C;Comment: The venom of this scorpion contains three major polypeptide neurotoxins: M9, C;Superfamily: scorpion neurotoxin

C;Keywords: neurotoxin; venom

F;12-65,16-36,22-46,26-48/Disulfide bonds: #status predicted

Alignment Scores:

Pred. No.:	0.418	Length:	66
Score:	76.00	Matches:	13
Percent Similarity:	59.62%	Conservative:	4
Best Local Similarity:	44.83%	Mismatches:	8
Query Match:	28.46%	Indels:	4
DB:	1	Gaps:	2

US-09-673-274B-2\_COPY\_1\_132 (1-132) x NTSR4E (1-66)

Qy 52 TGCACGGTGTGAGTGCAGAGGAGGGGTTACAGGGTGGTCACTGC-----GGTCC 102

Db 22 CysAspSerGluCysLysLysAsnGlyAlaAspGlySerTyrCysGlnTrpLeuGlyArg 41

Qy 103 TTGCTAAACGTGAACCTGCTGGTGGCGAG 129

Db 42 PheGlyAsnAla---CysTrpCysCys 49

RESULT 9

NTSR1C

neurotoxin 1 - bark scorpion (tentative sequence)

C;Species: Centruroides sculpturatus (bark scorpion)

C;Date: 22-Jun-1981 #sequence\_revision 22-Jun-1981 #text\_change 31-Mar-2000

C;Accession: A01752

R;Babin, D.R.; Watt, D.D.; Goos, S.M.; Mlejnek, R.V.

R;Arch. Biochem. Biophys. 164, 694-706, 1974

16 TCCTGCGTGTGGGTGTGTGAACACTACACTCCGATTCGACGGTGAAGAGGAGG 75

30 AsnCysValTyrHisCysVal-----ProProCysAspGlyLeuCysLysLysAsn 46

76 GGTTCACAAAGGTGTCACTCGGTCTCTCTC-----GCTAACGTGAACCTGCTGGTGC 126

47 GlyAlaLysSerGlySerCysGlyPheLeuIleProSerGlyLeuAlaCysTrpCys 65

RESULT 6

NTSR4L

neurotoxin IV - Egyptian scorpion

C;Species: Leiurus quinquestriatus (Egyptian scorpion)

C;Date: 15-Oct-1982 #sequence\_revision 27-Nov-1985 #text\_change 28-Aug-1998

C;Accession: A01749

R;Kopeyan, C.; Martinez, G.; Rochat, H.

R;BS Lett. 181, 211-217, 1985

A;Title: Primary structure of toxin IV of Leiurus quinquestriatus quinquestriatus.

A;Reference number: A01749

A;Accession: A01749

A;Molecule type: protein

A;Residues: 1-65 <KOP>

Note: the source is designated as Leiurus quinquestriatus quinquestriatus

C;Superfamily: scorpion neurotoxin

C;Keywords: blocked carboxyl end; neurotoxin; sodium channel inhibitor; venom

F;13-64,17-37,23-47,27-49/Disulfide bonds: #status predicted

F;65/Modified site: blocked carboxyl end (Arg) (probably amidated) #status experimental

Alignment Scores:

Pred. No.:	0.274	Length:	65
Score:	77.50	Matches:	16
Percent Similarity:	57.14%	Conservative:	8
Best Local Similarity:	38.10%	Mismatches:	9
Query Match:	29.03%	Indels:	9
DB:	1	Gaps:	4

16 TCCTGCGTGTGG-----GGTGTGTGAACACTACACTCCGATTCGACGGTGAAGAGGAGG 69

12 AsnCysValTyrThrCysGlySerAsnSerTyr-----CysAsnThrGluCysThr 28

70 AGGAGGGGTACAAAGGGTGTCACTGC-----GGTTCCTTCGCTAACGTGAAGTGC 120

29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpLeuGlyLysTyrGlyAsnAla---Cys 47

121 TGGTGC 126

48 TrpCys 49

RESULT 7

NTSR4L

neurotoxin I - scorpion (Leiurus quinquestriatus)

C;Species: Leiurus quinquestriatus quinquestriatus

C;Date: 06-Jul-1990 #sequence\_revision 06-Jul-1990 #text\_change 02-Sep-2000

C;Accession: B34123; S08268; B38526

R;Zlotkin, E.; Fowler, E.; Eitan, M.; Moyer, M.; Adams, M.E.

R;Toxin 28, 170, 1990

A;Title: On the chemistry and action of the depressant insect toxins.

A;Reference number: A34123

A;Accession: B34123

A;Molecule type: protein

A;Residues: 1-61 <ZLO>

R;Kopeyan, C.; Mansuette, P.; Sampieri, F.; Brando, T.; Bahraoui, E.M.; Rochat, H.; Grat

R;BS Lett. 261, 423-426, 1990

A;Title: Primary structure of scorpion anti-insect toxins isolated from the venom of Lei

R;Reference number: S08267; MUID:90184494; PMID:2311768

C;Accession: S08268

A;Molecule type: protein

A;Residues: 1-61 <KOP>

R;Zlotkin, E.; Eitan, M.; Bindokas, V.P.; Adams, M.E.; Moyer, M.; Burkhardt, W.; Fowler,

A:Title: Amino acid sequences of neurotoxic protein variants from the venom of Centruroides

A:Reference number: A9058; MUID:75163395; PMID:4460885

A:Accession: A01752

A:Molecule type: protein

A:Residues: 1-65 <BAB>

C:Superfamily: scorpion neurotoxin

C:Keywords: neurotoxin; venom

F:12-64,16-41,27-46,29-48/Disulfide bonds: #status predicted

Alignment Scores:

Pred. No.:	0.483	Length:	65
Score:	75.50	Matches:	14
Percent Similarity:	51.35%	Conservative:	5
Best Local Similarity:	37.84%	Mismatches:	15
Query Match:	28.28%	Indels:	3
DB:	1	Gaps:	1

US-09-673-274B-2\_COPY\_1\_132 (1-132) x NTSR:C (1-65)

QY 19 TCGTGTGGGTGCTGTGAACACTTCCGATTCGAACGGTGAAGTGAAGAGGAGG 78

DB 16 CysPreTrpLeuGlyLysAsnGluHisAsnTrpCysGluCysLysAlaLysAsnGlnGly 35

QY 79 TACAAGGGTGTCACTGGCGTTCCTTCGCTAAGCTGAAGTGTGTGGAG 129

DB 36 GlySerTrpGlyTyrCysTrpAlaPheAla-----CysTrpCysGlu 49

RESULT 10

A55824

Drosocypin precursor - fruit fly (*Drosophila melanogaster*)

N:Alternate names: cysteine-rich peptide

C:Species: *Drosophila melanogaster*

C:Date: 23-Mar-1995 #sequence\_revision 05 Apr 1995 #text\_change 21-Jul-2000

C:Accession: A55824; S38575

R:Rehbohm, P.; Bulet, P.; Michaut, L.; Lagieux, M.; Broekaert, W.F.; Hetru, C.; Hoffman

J.; Biol. Chem. 269, 33159-33163, 1994

A:Title: Insect immunity. Septic injury of *Drosophila* induces the synthesis of a potent

A:Reference number: A55824; MUID:95105209; PMID:7806546

A:Accession: A55824

A:Molecule type: mRNA

A:Residues: 1-70 <FEH>

A:Cross references: EMBL:X75595; NID:G414661; PIDD:CA53267.1; PID:G414662

A:Note: The mature peptide sequence was confirmed by peptide sequencing

C:Comment: The exact site of signal peptide cleavage (and the requirement, if any, for a

C:Genetics:

A:Gene: FlyBase: Drs

A:Cross references: FlyBase: FBgn0010381

C:Keywords: disulfide bond; extracellular protein

F:27-70/Product: drosocypin #status experimental <NAT>

Alignment Scores:

Pred. No.:	0.632	Length:	70
Score:	74.50	Matches:	15
Percent Similarity:	42.50%	Conservative:	2
Best Local Similarity:	37.50%	Mismatches:	16
Query Match:	27.90%	Indels:	7
DB:	2	Gaps:	3

US-09-673-274B-2\_COPY\_1\_132 (1-132) x A55824 (1-70)

QY 13 GTTCTCTGCG---GTGTGGGGTGTGTGAACACTTCCGATTCGAACGGTGAAGTGAAG 69

DB 35 GlyProCysAlaValTrpAsp-----AsnGluTrpCysArgValCysLys 50

QY 70 AGGAGGGGTACAAGGGTGTCACTGGCGTTCCTTCGCTAAGCTGAAGTGTGTGGAG 129

DB 51 GluGluGlyArgSerSerGlyHisCysSer-----ProSerLeuLysCysTrpCysGlu 68

RESULT 11

C23727

neurotoxin V-5 - bark scorpion

C:Species: *Centruroides sculpturatus* (bark scorpion)

C:Date: 31-Jan-1992 #sequence\_revision 31-Jan-1992 #text\_change 07-Feb-1997

C:Accession: C23727

R:David, R.M.; Krishna, N.R.; Watt, D.C.

Toxin 29, 645-662, 1991

A:Title: Characterization of cationic binding sites of neurotoxins from venom of the

A:Reference number: A23727; MUID:92023284; PMID:1926166

A:Accession: C23727

A:Molecule type: protein

A:Residues: 1-59 <DAV>

C:Superfamily: scorpion neurotoxin

C:Keywords: neurotoxin

Alignment Scores:

Pred. No.:	0.751	Length:	59
Score:	74.00	Matches:	16
Percent Similarity:	60.53%	Conservative:	7
Best Local Similarity:	42.11%	Mismatches:	7
Query Match:	27.72%	Indels:	8
DB:	2	Gaps:	3

US-09-673-274B-2\_COPY\_1\_132 (1-132) x C23727 (1-59)

QY 16 TCGTGTGGGTGCTGTGAACACTTCCGATTCGAACGGTGAAGTGAAGAGGAG 75

DB 14 SerCysVal-----AlaAsnAsnTrp-----CysAspAsnGlnCysLysMetLys 28

QY 76 GGTACAAAGGTGTCACTGGCGTTCCTTCGCTAAGCTGAAGTGTGTGGAG 129

DB 29 LysAlaSerGlyGlyHisCysTyrAla-----MetSerCysTyrCysGlu 43

RESULT 12

JC1321

neurotoxin IV - Sahara scorpion

C:Species: *Androctonus australis* (Sahara scorpion)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999

C:Accession: JC1321

R:Mansuelle, P.; Martin, M.F.; Rochat, H.; Granier, C.

Natural toxins 1, 61-69, 1992

A:Title: The amino acid sequence of toxin IV from the *Androctonus australis* scorpion:

A:Reference number: JC1321; MUID:94221398; PMID:1344902

A:Accession: JC1321

A:Molecule type: protein

A:Residues: 1-64 <MAN>

C:Superfamily: scorpion neurotoxin

C:Keywords: neurotoxin

Alignment Scores:

Pred. No.:	0.74	Length:	64
Score:	74.00	Matches:	12
Percent Similarity:	59.26%	Conservative:	4
Best Local Similarity:	44.44%	Mismatches:	9
Query Match:	27.72%	Indels:	2
DB:	2	Gaps:	1

US-09-673-274B-2\_COPY\_1\_132 (1-132) x JC1321 (1-64)

QY 52 TCGAACGGTGAAGTGAACAGAGGGGTACAGGGTGTCAAGGGTGTCAAGGGTCTTC 105

DB 20 CysAspGlyLeuCysLysLysAsnGlyAlaLysSerGlySerCysGlyPheLeuValPro 39

QY 106 GCTAACGTGAACCTGCTGGTGC 126

DB 40 SerGlyLeuAlaCysTrpCys 46

RESULT 13

NTSR5L

neurotoxin V - Egyptian scorpion

C:Species: *Leiurus quinquestriatus* (Egyptian scorpion)

C:Date: 30-Apr-1979 #sequence\_revision 13-Jul-1981 #text\_change 15-Oct-1996

C:Accession: A0174

R:Kopeyan, C.; Martinez, G.; Rochat, H.

FEBS Lett. 89, 54-58, 1978

A:Title: Amino acid sequence of neurotoxin V from the scorpion *Leiurus quinquestriatus*

A:Reference number: A01741; MUID:78191225; PMID:658402  
A:Note: L. q. quinquestratus

A:Accession: A01741

A:Molecule type: protein

A:Residues: 1-64 <KOP>

C:Superfamily: scorpion neurotoxin

C:Keywords: blocked carboxyl end; neurotoxin; venom

F:12-63,16-36,22-46,28-48/Disulfide bonds: #status predicted  
F:7,12-63,16-36,22-46,28-48/Disulfide bonds: #status experimental

F:7,64/Modified site: blocked carboxyl end (Asn) (probably amidated) #status experimental

Alignment Scores:

Pred. No.:	0.981	Length:	64
Score:	73.00	Matches:	12
Percent Similarity:	60.71%	Conservative:	5
Best Local Similarity:	42.86%	Mismatches:	7
Query Match:	27.34%	Indels:	4
DB:	1	Gaps:	2

US-09-673-274B-2\_COPY\_1\_132 (1-132) x NTSR5L (1-64)

2Y 52 TGCACGGTGTGACGACAGAGAGGGTGTACAGGGTGTCTACTGC-----GGTTC 102

Db 22 CysAsnAspGluCysLeuLysGlyGlySerGlyTyrCysGlnTrpAlaSerPro 41

2Y 103 TTCGCTACCGTGAACCTGCTGGTGC 126

Db 42 TyrGlyAsnAla---Cys:rpCys 49

RESULT 14

NTSR3C

C:Species: Centruroides sculpturatus (bark scorpion)

C:Date: 15-Oct-1982 #sequence\_revision 15-Oct-1982 #text\_change 23-Aug-1996

C:Accession: A90058; A94470; A01754

R:Rabin, D.R.; Watt, D.D.; Goos, S.M.; Mlejnek, P.V.

A:Title: Amino acid sequences of neurotoxic protein variants from the venom of Centruroides

A:Reference number: A90058; MUID:75163395; PMID:4460885

A:Accession: A90058

A:Molecule type: protein

A:Residues: 1-24, 'NTC', 28-63, 'CS' <BAB>

R:Rabin, D.R.; Watt, D.D.; Goos, S.M.; Mlejnek, P.V.

A:Note: unpublished results, cited by Fontecilla-Camps, J.C., et al., Toxicon 20, 1-7, 1982

A:Reference number: A94470

A:Accession: A94470

A:Molecule type: protein

A:Residues: 1-65 <BHO>

R:Fontecilla-Camps, J.C.; Almasay, R.J.; Suddath, F.L.; Bugg, C.E.

Toxicon 20, 1-7, 1982

A:Title: The three-dimensional structure of scorpion neurotoxins.

A:Reference number: A94314; MUID:82200153; PMID:7080025

A:Contents: annotation; X-ray crystallography, 1.8 angstroms; disulfide bonds

A:Note: X-ray crystallographic studies were based on the revised sequence shown

C:Superfamily: scorpion neurotoxin

C:Keywords: neurotoxin; venom

F:12-65,16-41,25-46,29-48/Disulfide bonds: #status experimental

Alignment Scores:

Pred. No.:	1.98	Length:	65
Score:	70.50	Matches:	15
Percent Similarity:	51.22%	Conservative:	6
Best Local Similarity:	36.59%	Mismatches:	13
Query Match:	26.40%	Indels:	7
DB:	1	Gaps:	3

US-09-673-274B-2\_COPY\_1\_132 (1-132) x NTSR3C (1-65)

2Y 19 TCGCTGTGGGTGCTGTGAACCTAC-----ACTTCGATTGCAACGGTGTGACGACAGAGG 72

Db 12 CysLeuTyrGlyCysLeuLysLeuGlyGluAsnGluGlyCysAspThrGluCysLysAla 31

2Y 73 AGGGGTACAGGGT-----CGTCACTGCGGTTCCTTCGCTAACGTCACCTGCTGGTGC 126

Db 73 AGGGGTACAGGGT-----CGTCACTGCGGTTCCTTCGCTAACGTCACCTGCTGGTGC 126

Db 32 LysAsnGlnGlyGlySerTyrGlyTyrCysTyrAlaPheAla-----CysTrpCys 48

QY 127 GAG 129

Db 49 Glu 49

RESULT 15

JN0670

Na+-channel-blocking toxin (clone cngtIII) precursor - scorpion (Centruroides nexius)

C:Species: Centruroides nexius

C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 16-Jul-1999

C:Accession: JN0670

R:Beckerl, B.; Vazquez, A.; Garcia, C.; Corona, M.; Bolivar, F.; Possani, L.D.

Gene 129, 165-171, 1993

A:Title: Cloning and characterization of cDNAs that code for Na+-channel blocking toxin

A:Reference number: JN0669; MUID:9322493; PMID:8390386

A:Accession: JN0670

A:Molecule type: mRNA

A:Residues: 1-87 <BEC>

A:Cross-references: GB:L05061; NID:9304569; FIDN:AAA28286.1; PID:q304569

A:Experimental source: venom gland

C:Superfamily: scorpion neurotoxin

C:Keywords: toxin

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-87/Product: Na+-channel-blocking toxin (clone cngtIII) #status predicted <VAT>

Alignment Scores:

Pred. No.:	1.88	Length:	87
Score:	70.50	Matches:	15
Percent Similarity:	48.72%	Conservative:	4
Best Local Similarity:	38.46%	Mismatches:	13
Query Match:	26.40%	Indels:	7
DB:	2	Gaps:	3

US-09-673-274B-2\_COPY\_1\_132 (1-132) x JN0670 (1-87)

QY 19 TCGCTGTGGGTGCTGTGAACCTACACTTCGATTGCAACGGTGTGACGACAGAGGACGCT 78

Db 35 CysPheTrpLeuGlyLysAsn-----GluGlyCysAspLysGluCysLysAlaLysAsn 52

QY 79 TACAAGGT-----SGTCACCTGCGGTTCCTTCGCTAACGTGAACCTGCTGGTCCGAG 129

Db 53 GlnGlyGlySerTyrGlyTyrCysTyrAlaPhe -----GlyCysTrpCysGlu 68

Search completed: October 16, 2003, 17:11:59

Cob time : 17.2259 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

MM nucleic - protein search, using frame\_plus\_n2p model  
Run on: October 16, 2003, 17:06:37 ; Search time 39.0299 Seconds  
(without alignments)  
1745.475 Million cell updates/sec

Title: US-09-673-274B-2\_COPY\_1\_132

Perfect score: 267

Sequence: 1 gataagcttcaggcttcg.....tgaactgctggtcgagact 132

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0

Maximum Match 100%

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

MODEL=frame+n2p.model DEV=xlp  
Q=/cgn2/1/USPRO-epool/US09673274/runat\_16102003\_170627\_2545/app\_query.fasta\_1.654  
DB=SPTREMBL\_23 -QFMT=fastan -SUFFIX=tspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
USER=US09673274 -CGEN 1\_216 -runat\_16102003\_170627\_2545 -NCPU=6 -ICPU=3  
NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL\_23.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_plant.\*  
10: sp\_protist.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84.5	31.6	65	5 Q17231	Q17231 mesobuthus

ALIGNMENTS

RESULT 1

Q17231  
ID Q:7231 PRELIMINARY, PRT: 65 AA.  
AC Q17231  
DT 01-NOV-1996 (TREMBL:rel. 01, Created)  
DT 01-NOV-1996 (TREMBL:rel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMBL:rel. 23, Last annotation update)  
DE Insect toxin 2.  
GN IT2.

OS Mesobuthus martensii (Manchurian scorpion) (Buttus martensii).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthoidea; Buthidae; Mesobuthus.  
CX NCBI\_TaxID=34649;  
RN [1]  
PP SEQUENCE FROM N.A.  
RC TISSUE=telson;  
RA Zhu X., Zhang T., Zhu Y.;

RT "Cloning and sequencing of two depressant insect selective neurotoxin cDNAs from Buttus martensii Karsch."  
RL Kexue Tongbao 41:1387-139 (1996)  
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
DR EMBL; X92077; CAA63059.1; -  
DR HSSP; P01494; 2SN3.  
DR InterPro; IPR003614; Knot1.

Q17230 mesobuthus  
Q810K7 mesobuthus  
Q95WX6 mesobuthus  
Q9XY87 mesobuthus  
Q8T3C0 mesobuthus  
P82761 arabidopsis  
Q9VZQ5 drosophila  
Q8MVS7 leirus qui  
Q9YU3 mesobuthus  
Q9VZ2 drosophila  
Q8WY4 centruroides  
Q9VZ3 drosophila  
Q9VZ6 mesobuthus  
Q9VZ9 mesobuthus  
Q9VZ4 drosophila  
Q9Y0B8 mesobuthus  
Q39403 brassica ca  
Q8VHS2 mus musculus  
Q9TW11 centruroides  
P91904 caenorhabdi  
Q76963 orthochirus  
Q99H52 nectria rad  
Q9BIM4 androctonus  
Q9WHF2 agrotis seg  
Q26292 leirus qui  
P82759 arabidopsis  
Q8BMD9 mus musculus  
Q8AXP0 cyrops pyr  
Q9BIMC androctonus  
P82762 arabidopsis  
Q8IRD7 drosophila  
Q9BIM2 androctonus  
Q945D8 stratiococcus  
Q9VZQ0 streptococcus  
Q9VZQ0 drosophila  
Q4367 zea mays (m  
Q8NHD4 homo sapien  
Q8WY3 homo sapien





FT SIGNAL 1 21 POTENTIAL.  
CHAIN 22 82 DEPRESSANT INSECT NEUROTOXIN BMK ITB.  
SEQUENCE 85 AA; 9330 MW; 80CDD3914956D8C CRC64;

Alignment Scores:  
Pred. No.: 0.0268 Length: 85  
Score: 81.50 Matches: 15  
Percent Similarity: 55.26% Conservative: 6  
Best Local Similarity: 39.47% Mismatches: 12  
Query Match: 30.52% Indels: 5  
DB: Gaps: 2

US-09-673-274B-2\_COPY\_1\_132 (1-132) x Q95WX6 (1-85)

QY 16 TCCTCGGTGGGTGGTGGTGAACCTCCGATTGCAACGGTGAGTGCAGAGGAGG 75  
|||||:|||||  
Db 34 SerCysLeuTrpGly-----AsnGluGlyCysAsnLysGluCysLysAlaPhe 49  
|||||:|||||

QY 76 GGTTCACAGGGTGGTCACTCGGTTCTTCGCTAACGTTGAACGTTGGTGGCGAG 129  
|||||:|||||  
Db 50 GlyAlaSerTyGlyTyrcysTrpThrTrp---GlyLeuAlaCysTrpCysGln 66  
|||||:|||||

RESULT 5  
Q95WX6  
ID Q95X87 PRELIMINARY; PRT; 85 AA.  
AC Q95X87  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
CC -|- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).  
CC -|- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
DR EMBL: AF064821; AAD31592.1; -  
DR HSSP: P01434; 2SN3.  
DR InterPro: IPR003614; Knot1.  
DR Pfam: PF00537; toxin\_3; 1.  
DR ProDom: PD00908; Scorpion\_toxinL; 1.  
DR SMART: SM00505; Knot1; 1.  
DR Neurotoxin: Signal.  
DR SIGNAL 1 21 POTENTIAL.  
DR CHAIN 22 85 NEUROTOXIN AEP.  
DB 50 GlyAlaSerTyGlyTyrcysTrpThrTrp---GlyLeuAlaCysTrpCysGln 66  
|||||:|||||

Alignment Scores:  
Pred. No.: 0.0268 Length: 85  
Score: 81.50 Matches: 14  
Percent Similarity: 55.26% Conservative: 7  
Best Local Similarity: 36.84% Mismatches: 12  
Query Match: 30.52% Indels: 5  
DB: Gaps: 2

US-09-673-274B-2\_COPY\_1\_132 (1-132) x Q95X87 (1-85)

QY 16 TCCTCGGTGGGTGGTGGTGAACCTCCGATTGCAACGGTGAGTGCAGAGGAGG 75  
|||||:|||||  
Db 34 SerCysLeuTrpGly-----AsnGluGlyCysAsnLysGluCysLysAlaPhe 49  
|||||:|||||

QY 76 GGTTCACAGGGTGGTCACTCGGTTCTTCGCTAACGTTGAACGTTGGTGGCGAG 129  
|||||:|||||  
Db 50 GlyAlaSerTyGlyTyrcysTrpThrTrp---GlyLeuAlaCysTrpCysGln 66  
|||||:|||||

RESULT 6  
Q95X87  
ID Q95X87 PRELIMINARY; PRT; 85 AA.  
AC Q95X87  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
CC -|- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).  
CC -|- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
DR EMBL: AF064821; AAD31592.1; -  
DR HSSP: P01434; 2SN3.  
DR InterPro: IPR003614; Knot1.  
DR Pfam: PF00537; toxin\_3; 1.  
DR ProDom: PD00908; Scorpion\_toxinL; 1.  
DR SMART: SM00505; Knot1; 1.  
DR Neurotoxin: Signal.  
DR SIGNAL 1 21 POTENTIAL.  
DR CHAIN 22 85 NEUROTOXIN AEP.  
DB 50 GlyAlaSerTyGlyTyrcysTrpThrTrp---GlyLeuAlaCysTrpCysGln 66  
|||||:|||||

Q8T3T0;  
AC 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Depressant insect toxin BMK 11al precursor.  
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthoidea; Buthidae; Mesobuthus.  
OX NCBI\_TaxID=34649;  
RN [-]  
RP SEQUENCE FROM N.A.  
RA Liu Z., Chi C., Wu X.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
CC -|- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).  
CC -|- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
DR EMBL: AY090782; AA090937.1; -  
DR InterPro: IPR003614; Knot1.  
DR Pfam: PF00537; toxin\_3; 1.  
DR ProDom: PD00908; Scorpion\_toxinL; 1.  
DR SMART: SM00505; Knot1; 1.  
DR Signal.  
FT SIGNAL 1 21 POTENTIAL.  
SQ SEQUENCE 85 AA; 9214 MW; 6CF567CB60323C1A CRC64;

Alignment Scores:  
Pred. No.: 0.0366 Length: 85  
Score: 80.50 Matches: 15  
Percent Similarity: 52.63% Conservative: 5  
Best Local Similarity: 39.47% Mismatches: 13  
Query Match: 30.15% Indels: 5  
DB: Gaps: 2

US-09-673-274B-2\_COPY\_1\_132 (1-132) x Q8T3T0 (1-85)

QY 16 TCCTCGGTGGGTGGTGGTGAACCTCCGATTGCAACGGTGAGTGCAGAGGAGG 75  
|||||:|||||  
Db 34 SerCysLeuTrpGly-----AsnGluGlyCysAsnLysGluCysLysAlaPhe 49  
|||||:|||||

QY 76 GGTTCACAGGGTGGTCACTCGGTTCTTCGCTAACGTTGAACGTTGGTGGCGAG 129  
|||||:|||||  
Db 50 GlyAlaSerTyGlyTyrcysTrpThrTrp---GlyLeuAlaCysTrpCysGln 66  
|||||:|||||

RESULT 7  
Q8T3T0  
ID P82761 PRELIMINARY; PRT; 91 AA.  
AC P82761  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein LCR46 precursor.  
GN LCR46.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1];  
RP SEQUENCE FROM N.A.  
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,  
RA Miyajima N., Tabata S.;  
RL "Structural analysis of Arabidopsis thaliana chromosome 5. IX. Sequence features of the regions of 1,011,550 bp covered by seventeen P1 and TAC clones.";  
RT DNA Res. 6:183-195(1999).  
RN [2];  
RP IDENTIFICATION.  
RX PubMed=11437247;  
RA Vanosthuyse V., Mieg C., Dumas C., Cock J.M.;  
RT "Two large Arabidopsis thaliana gene families are homologous to the Brassica gene superfamily that encodes pollen coat proteins and the

RT Mate component of the self-incompatibility response.";  
 KL Plant Mol. Biol. 46:17-34 (2001).  
 DR EMBL: AB017061; -; NOT ANNOTATED CDS.  
 KW Hypothetical protein; Signal; POTENTIAL.  
 FT SIGNAL 1 27  
 FT CHAIN 28 91  
 SQ SEQUENCE 91 AA; 9503 MW; 10BDEA8E8369F55 CRC64;

Alignment Scores:  
 Pred. No.: 0.0429 Length: 91  
 Score: 80.00 Matches: 14  
 Percent Similarity: 58.08% Conservative: 4  
 Best Local Similarity: 45.16% Mismatches: 23  
 Query Match: 29.96% Indels: 0  
 DB: 10 Gaps: 0

US-09-673-274B-2\_COPY\_1\_132 (1-132) x P8276: (1-91)

QY 16 TCTCGTGTCGGGCTGTGTAACCTTCGATTCGACGGTGAATGCAAGAGGAGG 75  
 |||||  
 LL 31 ThrCysLeuProGlyGluCysThrAsnProSerGluCysAsnAlaAlaCysLysSerAsn 50  
 |||||  
 QY 76 GGTACAGGTCGTCACATCGCGTTCCTCCCT 108  
 |||||  
 DD 5: GlyTyrLysGlyGlyAlaCysValSerMetSer 6:

## RESULT 8

Q9VZQ5 PRELIMINARY; PRT; 69 AA.  
 AC Q9VZQ5;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
 DT 01-MAY-2003 (TRENBLrel. 23, Last annotation update)  
 DE C332274 protein.  
 CN C332268 OR C332815 OR C331520.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydriidea; Drosophilidae; Drosophila  
 OX NCBI\_TaxID=7227;  
 RN (1-91)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Anandides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt A.D., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhardani D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherly J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos G., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier A., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glueck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Hartis N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Houston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaisi M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li C., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel B., McIntosh T.C., McLeod X.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarty C., Morris C., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.B.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Siripson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RC "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Anandides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Dou L.E., Doyle C., Dreanek D., Farfan D.,  
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattel B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Pacleeb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RC "Sequencing of Drosophila melanogaster genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Tu P.J., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungali C.J., Lewis S.E.;  
 RC "Annotation of Drosophila melanogaster genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN (5)  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003477; AAF47764.2; -;  
 DR HSSP: P41964; 1MW;  
 DR FlyBase: FBgn032268; CG32268.  
 DR InterPro: IPR002118; Gamma-tubulin.  
 DR ProDom: PD002594; Gamma-tubulin; 1.  
 SQ SEQUENCE 69 AA; 7720 MW; 87CC05D254A11AC4 CRC64;

Alignment Scores:  
 Pred. No.: C.0677 Length: 69  
 Score: 78.50 Matches: 16  
 Percent Similarity: 51.28% Conservative: 4  
 Best Local Similarity: 41.03% Mismatches: 14  
 Query Match: 29.40% Indels: 5  
 DB: 5 Gaps: 2

US-09-673-274B-2\_COPY\_1\_132 (1-132) x Q9VZQ5 (1-69)

QY 13 GGTTCTCGGTGGGGTGGTGTGTAACCTTCGATTCGACGGTGAAGAGG 72  
 |||||  
 Db 34 GlySerCys-----AlaValTprHisArgLysCysValAspIleCysGlnArg 50  
 |||||  
 QY 73 AGGGGTACAGGGTGGTGTGCTACTCGGTTCTCTGCTAACGTGAACCTCGTGGTGGAG 129  
 |||||  
 Db 51 GluGlyArgThrSerGlyHisCysSer-----ProSerLeuLysCysTrpCysGlu 67  
 |||||

## RESULT 9

Q8MV57 PRELIMINARY; PRT; 82 AA.  
 ID Q8MV57

```
Q8MVS7:
01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Depressant insect toxin 2.
Leirus quinquestratus quinquestratus (Egyptian scorpion).
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
Buthoidea; Buthidae; Leiurus.
NCBI_TaxID=6885;
[1]_SEQUENCE FROM N.A.
Zaki T.I., Marunak J.E.;
*Three polymorphic genes encoding Lqqt12 from the Egyptian scorpion
Leirus quinquestratus quinquestratus.
Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
-!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
-!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
EMBL; AF474984; AAM74028.1; -.
InterPro: IPR003614; Knott1.
InterPro: IPR002061; Scorpion_toxinL.
Pfam: PF00537; toxin_3; 1.
ProDom: PD000908; Scorpion_toxinL; 1.
SMART: SM00505; Knott1; 1.
SEQUENCE 82 AA; 9099 MW; F13D53B18CDECBF9 CRC64;

Alignment Scores:
Pred. No.: 0.127 Length: 82
Score: 76.50 Matches: 15
Percent Similarity: 55.26% Conservative: 6
Best Local Similarity: 39.47% Mismatches: 12
Query Match: 28.65% Indels: 5
Gaps: 2
DB:

US-09-673-274b-2_copy_1_132 (1-132) x Q8MVS7 (1-82)
QY 16 TCCTCGGTGGGTGGTGTGTAACACTCCGATTCGACGAGTGTGCAAGAGGAGG 75
D 16 TCCTCGGTGGGTGGTGTGTAACACTCCGATTCGACGAGTGTGCAAGAGGAGG 75
Db 34 SerCysLeuPheGly-----AsnGluGlyCysAsnLysGluCysLysSerTyr 49
QY 76 GGTACAGGTTGCTACGTCGGTTCCTTCGTAACGTAACGTAACGTCGTCGTCGAG 129
D 76 GGTACAGGTTGCTACGTCGGTTCCTTCGTAACGTAACGTAACGTCGTCGTCGAG 129
Db 50 GlyGlySerTyrGlyTyrCysrprhtrp---GlyLeuAlaCysrprCysGlu 66

RESULT 10
QYVZR2
ID QYVZR2 PRELIMINARY; PRT; 69 AA.
AC QYVZR2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created);
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update);
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update);
DE CG10812 protein (GH09576p);
GN CG10812.
OS Drosophila melanogaster (Fruit fly);
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]_SEQUENCE FROM N.A.
RC S-RAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.P., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Berens P.V., Bertram R.P., Bhandari D., Bolshakov E.M.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Rurtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew J., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evgelisita C.C., Ferraz C., Ferreira S., Fleischman K.,
RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaisi M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li C., Li Z., Liang Y., Lin X.,
RA Liu X., Mattie B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Neilson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb C.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Ye R.-F., Zaveri J.S., Zhan M., Zhou G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.C.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

PM nucleic - protein search, using frame\_plus\_n2p model

Run on: October 16, 2003, 17:12:07 : Search time 27.4086 Seconds  
(without alignments)  
1578.850 Million cell updates/sec

Title: US-09-673-274b-2\_COPY\_1\_132  
Perfect score: 267  
Sequence: 1 gataagcttaccggttcctg.....tgaactgctggtgcgagact 132

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 609560 seqs, 16391702 residues

Total number of hits satisfying chosen parameters: 1219120

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

MODEL=frame+2p.model -DEV=xlp  
Q=/cgn2\_1/USPRO\_pool/US09673274/runat\_16102003\_170630\_2743/app\_query\_fasta\_1.654  
DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1  
LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blcsum62  
TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100  
THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -M=NLEN=0  
MAXLEN=2000000000 -USER=US09673274@cgn\_1\_106@runat\_16102003\_170630\_2743  
-NCPU=3 -NO WMAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/2/pubaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

1	232	86.9	44	15	US-10-125-258-119
2	232	86.9	73	15	US-10-125-258-117
3	79.5	29.8	65	15	US-10-264-480-13
4	79.5	29.8	66	15	US-10-264-480-12
5	76.5	28.7	61	15	US-10-264-480-9
6	75	28.1	61	9	US-09-829-481-2
7	74.5	27.9	44	12	US-10-180-247-4
8	74.5	27.9	69	12	US-10-180-247-6
9	74.5	27.9	70	12	US-10-180-247-2
10	71.5	26.8	61	15	US-10-264-480-6
11	71.5	26.8	61	15	US-10-264-480-7
12	71.5	26.8	61	15	US-10-264-480-8
13	71	26.6	47	15	US-10-178-213-411
14	71	26.6	77	15	US-10-178-213-410
15	70.5	26.4	47	15	US-10-178-213-387
16	70.5	26.4	65	15	US-10-264-480-14
17	70.5	26.4	84	15	US-10-178-213-386
18	70.5	26.4	87	15	US-10-264-480-10
19	69.5	26.0	84	15	US-10-264-480-18
20	69	25.8	47	15	US-10-178-213-408
21	69	25.8	47	15	US-10-178-213-444
22	69	25.8	77	15	US-10-178-213-407
23	69	25.8	78	15	US-10-178-213-443
24	68.5	25.1	173	15	US-10-156-761-7787
25	68	25.5	47	15	US-10-178-213-396
26	68	25.5	76	15	US-10-178-213-395
27	67.5	25.3	87	15	US-10-264-480-12
28	67	25.1	45	15	US-10-178-213-168
29	67	25.1	72	15	US-10-178-213-167
30	67	25.1	86	15	US-10-264-480-22
31	67	25.1	86	15	US-10-264-480-23
32	66.5	24.9	47	15	US-10-178-213-309
33	66.5	24.9	47	15	US-10-178-213-339
34	66.5	24.9	58	15	US-10-264-480-4
35	66.5	24.9	62	14	US-10-044-359-21
36	66.5	24.9	66	15	US-10-264-480-15
37	66.5	24.9	77	15	US-10-178-213-308
38	66.5	24.9	84	15	US-10-178-213-338
39	65.5	24.5	44	15	US-10-178-213-147
40	65.5	24.5	58	15	US-10-264-480-3
41	65.5	24.5	71	15	US-10-178-213-146
42	65	24.3	45	15	US-10-178-213-210
43	65	24.3	47	15	US-10-178-213-3
44	65	24.3	47	15	US-10-178-213-457
45	65	24.3	47	15	US-10-178-213-450

## ALIGNMENTS

RESULT 1  
US-10-125-258-119  
: Sequence 119, Application US/10125258  
: Publication No. US20030028920A1  
: GENERAL INFORMATION:  
: APPLICANT: Altier, Daniel J.  
: APPLICANT: Herrana, Rafael  
: APPLICANT: L.J. Albert L.  
: APPLICANT: McCutchen, Billy F.  
: APPLICANT: Presnail, James K.  
: APPLICANT: Weaver, Janine L.  
: APPLICANT: Wong, James F. H.  
: TITLE OF INVENTION: Antimicrobial Polypeptides and Their  
: FILE OF INVENTION: Uses  
: FILE REFERENCE: 35718/246215  
: CURRENT APPLICATION NUMBER: US/10/125,258  
: CURRENT FILING DATE: 2002-04-18  
: PRIOR APPLICATION NUMBER: 60/285,355  
: PRIOR FILING DATE: 2001-04-20  
: NUMBER OF SEQ ID NOS: 127  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 119  
: LENGTH: 44  
: TYPE: PRT





```

: ORGANISM: Scolopendra canidens DS
US-09-829-481-2
Alignment Scores:
Pred. No.: 0.33 Length: 61
Score: 75.00 Matches: 11
Percent Similarity: 53.57% Conservative: 4
Best Local Similarity: 39.29% Mismatches: 13
Query Match: 28.09% Indels: 0
DB: Gaps: 0
US-09-673-274B-2_COPY_1_132 (1-132) x US-09-829-481-2 (1-61)
Qy 49 GATTGCAACGGTGAGTGCACAGAGGAGGGTTTACAGAGGGTGGTCACTCGGTTCTTGGCT 108
Db 34 GluCysAsnArgHisCysArgG.YAsnGlyPheThrGlyGlyCysThrGlyPheLeu 53
Qy 109 AACGTGAACGTGCTGGTGGAGACT 132
Db 54 LysPheThrCysLysCysTyrThr 61
RESULT 7
US-10-180-247-4
; Sequence 4, Application US/10180247
; Publication No. US2003016759A1
; GENERAL INFORMATION:
; APPLICANT: DeRose, Richard
; APPLICANT: Freysinet, Georges
; APPLICANT: Hoffman, Jules
; TITLE OF INVENTION: Chimeric Gene Encoding Drosomyoin,
; TITLE OF INVENTION: Vector Containing It and Production of Disease-Resistant
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: A32889-PCT-USA-A-A 072667.0192
; CURRENT APPLICATION NUMBER: US/10180247
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 09/480,251
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/FR98/01462
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: FR97/03,115
; PRIOR FILING DATE: 1997-07-11
; PRIOR APPLICATION NUMBER: FR97/09,663
; PRIOR FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-180-247-4
Alignment Scores:
Pred. No.: 6.371 Length: 44
Score: 74.50 Matches: 15
Percent Similarity: 42.50% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 16
Query Match: 27.90% Indels: 7
DB: Gaps: 3
US-09-673-274B-2_COPY_1_132 (1-132) x US-10-180-247-4 (1-44)
Qy 13 GGTTCTTCCTGC---GTGTGGGGTGTGTGAACACTACACTTCGATTCGACGGTGGTGAAG 69
Db 9 GlyProCysAlaValTrpAsp-----AsnGluThrCysArgValCysLys 24
Qy 70 AGGAGGGGTTACAGGGTGTGTCACTCGGTTCTTCCCTAACTGAACTGCTGCTGCGAG 129
Db 25 GluGluG.YArgSerSerGlyH.sCysSer-----ProSerLeuLysCysTrpCysGlu 42
RESULT 8
US-10-180-247-6
; Sequence 6, Application US/10180247
; Publication No. US20030167519A1
: ORGANISM: Scolopendra canidens DS
US-09-829-481-2
Alignment Scores:
Pred. No.: 0.212 Length: 61
Score: 76.50 Matches: 15
Percent Similarity: 55.26% Conservative: 6
Best Local Similarity: 39.47% Mismatches: 12
Query Match: 28.65% Indels: 5
DB: Gaps: 2
US-09-673-274B-2_COPY_1_132 (1-132) x US-10-264-480-9 (1-61)
Qy 16 TCCTGCGTGGTGGTGGTGTGTAACACTACACTTCGATTCGACGGTGGTGAAGAGGAGG 75
Db 13 SerCysLeuPheGly-----AsnGluGlyCysAsnLysGluCysLysSerTyr 28
Qy 76 GGTACAAAGGGTGGTCACTCGGTTCTTCGTAACGTGAACGTGGTGGTGGCAG 129
Db 29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45
RESULT 6
US-09-829-481-2
Sequence 2, Application US/09829481
Patent No. US20020069427A1
GENERAL INFORMATION:
APPLICANT: Presnail, James
APPLICANT: Weng, Zude
APPLICANT: Wong, James
TITLE OF INVENTION: Arthropod Defensins
FILE REFERENCE: BB1441 US NA
CURRENT APPLICATION NUMBER: US/09/829,481
CURRENT FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/197279
PRIOR FILING DATE: 2000-04-14
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Microsoft Office 97
SEQ ID NO 2
LENGTH: 61
TYPE: PRT
: ORGANISM: Scolopendra canidens DS
US-09-829-481-2
Alignment Scores:
Pred. No.: 0.212 Length: 61
Score: 76.50 Matches: 15
Percent Similarity: 55.26% Conservative: 6
Best Local Similarity: 39.47% Mismatches: 12
Query Match: 28.65% Indels: 5
DB: Gaps: 2
US-09-673-274B-2_COPY_1_132 (1-132) x US-10-264-480-9 (1-61)
Qy 16 TCCTGCGTGGTGGTGGTGTGTAACACTACACTTCGATTCGACGGTGGTGAAGAGGAGG 75
Db 13 SerCysLeuPheGly-----AsnGluGlyCysAsnLysGluCysLysSerTyr 28
Qy 76 GGTACAAAGGGTGGTCACTCGGTTCTTCGTAACGTGAACGTGGTGGTGGCAG 129
Db 29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45
RESULT 6
US-09-829-481-2
Sequence 2, Application US/09829481
Patent No. US20020069427A1
GENERAL INFORMATION:
APPLICANT: Presnail, James
APPLICANT: Weng, Zude
APPLICANT: Wong, James
TITLE OF INVENTION: Arthropod Defensins
FILE REFERENCE: BB1441 US NA
CURRENT APPLICATION NUMBER: US/09/829,481
CURRENT FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/197279
PRIOR FILING DATE: 2000-04-14
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Microsoft Office 97
SEQ ID NO 2
LENGTH: 61
TYPE: PRT

```

```

1 GENERAL INFORMATION:
2 APPLICANT: Derose, Richard
3 APPLICANT: Freysinnet, Georges
4 APPLICANT: Hoffman, Jules
5 TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
6 VECTOR CONTAINING IT AND PRODUCTION OF DISEASE-RESISTANT
7 PLANTS
8 TITLE OF INVENTION: Transgenic Plants
9 FILE REFERENCE: A32889-PCT-USA-A-A 072667.0182
10 CURRENT APPLICATION NUMBER: US/1C/180,247
11 CURRENT FILING DATE: 2002-06-26
12 PRIOR APPLICATION NUMBER: 09/480,251
13 PRIOR FILING DATE: 2000-01-11
14 PRIOR APPLICATION NUMBER: PCT/FR98/01462
15 PRIOR FILING DATE: 1998-07-08
16 PRIOR APPLICATION NUMBER: FR97/09,115
17 PRIOR FILING DATE: 1997-07-11
18 PRIOR APPLICATION NUMBER: FR97/09,663
19 PRIOR FILING DATE: 1997-07-24
20 NUMBER OF SEQ ID NOS: 15
21 SOFTWARE: FastSeq for Windows Version 4.0
22 SEQ ID NO 6
23 LENGTH: 69
24 TYPE: PRT
25 ORGANISM: Artificial Sequence
26 FEATURE:
27 OTHER INFORMATION: Ubiquitin drosomycin fusion peptide
28 US 10 180 247 6
29
30 Alignment Scores:
31 Pctd. No.: 0.387 Length: 69
32 Score: 74.50 Matches: 15
33 Percent Similarity: 42.50% Conservative: 2
34 Best Local Similarity: 37.50% Mismatches: 16
35 Query Match: 27.90% Indels: 7
36 DS: 12 Gaps: 3
37
38 US-09-673-274B-2_COPY_1_132 (1-132) x US-10-180 247-6 (1-69)
39
40 QY 13 GGTCCTCG---GTGGGCTGCTGAACACTCACTTCCAGTGCACSGTGAGTGCAG 69
41 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
42 DB 14 GlyProCysAlaValTrpAsp-----AacGUThrCysArgValCysLys 49
43
44 QY 70 AGAGGGGTTCACAAGGGTGGTCACTGCCGTTCTTCGTACAGTGAATCGTGGTGGCAG 129
45 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
46 DB 50 GIUGLUGIyArgSerSetGlyHisCysSe-----ProSerLeuLysCysTrpCysGlu 67
47
48 RESULT 9
49 US-10-180-247-2
50 Sequence 2, Application US/1018C247
51 Publication No. US20030167519A1
52 GENERAL INFORMATION:
53 APPLICANT: Derose, Richard
54 APPLICANT: Freysinnet, Georges
55 APPLICANT: Hoffman, Jules
56 TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
57 VECTOR CONTAINING IT AND PRODUCTION OF DISEASE-RESISTANT
58 PLANTS
59 TITLE OF INVENTION: Transgenic Plants
60 FILE REFERENCE: A32889-PCT-USA-A-A 072667.0182
61 CURRENT APPLICATION NUMBER: US/1C/180,247
62 CURRENT FILING DATE: 2002-06-26
63 PRIOR APPLICATION NUMBER: 09/480,251
64 PRIOR FILING DATE: 2000-01-11
65 PRIOR APPLICATION NUMBER: PCT/FR98/01462
66 PRIOR FILING DATE: 1998-07-08
67 PRIOR APPLICATION NUMBER: FR97/09,115
68 PRIOR FILING DATE: 1997-07-11
69 PRIOR APPLICATION NUMBER: FR97/09,663
70 PRIOR FILING DATE: 1997-07-24
71 NUMBER OF SEQ ID NOS: 15
72 SOFTWARE: FastSeq for Windows Version 4.0
73 SEQ ID NO 2
74 LENGTH: 70
75 TYPE: PRT

```

```

; ORGANISM: Drosophila melanogaster
US-10-180-247-2

Alignment Scores:
Pred. No.:      0.388          Length:      70
Score:         74.50        Matches:       15
Percent Similarity: 42.50%    Conservative: 2
Best Local Similarity: 37.50% Mismatches:   16
Query Match:     27.90%     Indels:       7
DB:              12         Gaps:        3

US-09-673-274B-2_COPY_1_132 (-132) x US-10-180-247-2 (1-70)
QY      13 GGTCTCTGCG---GTGTGGGGTGTGAACATACACTCCGATTGCAACGGTGAGTGCAGG 199
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      35 GLYPFCYSKAAVALTPASP-----AsnGLUThrCysArgArGValCysLys 150
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY      70 AGCAGGGGTTCACAAGGGTGCATCGCGTTCTTCCTTAACGTGAACACTGCTGGTGGAG 129
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      51 GIUG-uGIyAqgSerSerGlyHisCysSer-----ProSerLeuLysCysTrpCysGlu 68
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-10-264-480-6
; Sequence 6, Application US/10264480
; Publication NO. US20030113892A1
; GENERAL INFORMATION:
; APPLICANT: Hammock, Bruce D.
; TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
; TITLE OF INVENTION: FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
; FILE REFERENCE: UCAL256
; CURRENT APPLICATION NUMBER: US/10/264,480
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/393,070
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/327,602
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Buthacus arenicola
US-10-264-480-6

Alignment Scores:
Pred. No.:      0.932          Length:      61
Score:         71.50        Matches:       14
Percent Similarity: 55.26%    Conservative: 7
Best Local Similarity: 36.84% Mismatches:   12
Query Match:     26.78%     Indels:       5
DB:              15         Gaps:        2

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-264-480-6 (1-61)
QY      16 TCTGCGTGTGGGGTGC*GTGAACATACACTCCGATTGCAACGGTGAGTGCAGGAGG 75
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      13 SerCysLeuPheGly-----AsnGLUGLyCysAspLysGluCysLysAlaItyr 29
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY      76 GGTTCACAAGGGTGCATCGCGTTCTTCCTTAACGTGAACACTGCTGGTGGAG 129
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      29 GlyGlySerTy-GlyTyrcyslrprthrpr---GlyLeuAlaCysTrpCysGlu 45
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-10-264-480-7
; Sequence 7, Application US/10264480
; Publication NO. US20030113892A1
; GENERAL INFORMATION:
; APPLICANT: Hammock, Bruce D.
; TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
; TITLE OF INVENTION: FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
; FILE REFERENCE: UCAL256

```

```

; ORGANISM: Drosophila melanogaster
US-10-180-247-2

Alignment Scores:
Pred. No.:      0.388          Length:      70
Score:         74.50          Matches:     15
Percent Similarity: 42.50%       Conservative: 2
Best Local Similarity: 37.50%    Mismatches:  16
Query Match:   27.90%           Indels:      7
DB:            1.2             Gaps:        3

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-180-247-2 (1-70)
QY      13 GGTCTCTGC---GTTGGGGTGTGAACATACACTCCGATTGCAACGGTGAGTCAAG 69
Db      35 GlyPrcCysA.aValTrpAsp-----AsnGluThrCysArgArGValCysLys 50
CY      70 AGCAGGGGTTCACAAGGGTGTCATCGCGTTCCTTCGCTAACGTGAACACTGCTGGTGGAG 129
Db      51 GluG-uGlyArgSerGlyHisCysSer-----ProSerLeuLysCysTrpCysGlu 68

RESULT 10
US-10-264-480-6
; Sequence 6, Application US/10264480
; Publication NO. US20030113892A1
; GENERAL INFORMATION:
; APPLICANT: Hammock, Bruce D.
; TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
; TITLE OF INVENTION: FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
; FILE REFERENCE: UCAL256
; CURRENT APPLICATION NUMBER: US/10/264,480
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/393,070
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/327,602
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 61
; TYPE: PRF
; ORGANISM: Buthacus arenicola
US-10-264-480-6

Alignment Scores:
Pred. No.:      0.932          Length:      61
Score:         71.50          Matches:     14
Percent Similarity: 55.26%       Conservative: 7
Best Local Similarity: 36.84%    Mismatches:  12
Query Match:   26.78%           Indels:      5
DB:            1.5             Gaps:        2

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-264-480-6 (1-61)
QY      16 TCTCGCTGTGGGGTGC*GTGAACATACACTCCGATTGCAACGGTGAGTCAAGCAGG 75
Db      13 SerCysLeuPhcGly-----AsnGluGlyCysAspLysGluCysLysAlaI Tyr 28
CY      76 GGTTCACAAGGGTGCTAACTGCGCGTTCCTTCGCTAACGTGAACACTGCTGGTGGAG 129
Db      29 GlyGlySerTy-GlyTyrcysrprthrpr---GlyLeuAlaCysTrpCysGlu 45

RESULT 11
US-10-264-480-7
; Sequence 7, Application US/10264480
; Publication NO. US20030113892A1
; GENERAL INFORMATION:
; APPLICANT: Hammock, Bruce D.
; TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS
; TITLE OF INVENTION: FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE
; FILE REFERENCE: UCAL256

```

CURRENT APPLICATION NUMBER: US/10/264,480  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: 60/393,070  
PRIOR FILING DATE: 2002-06-28  
PRIOR APPLICATION NUMBER: 60/327,602  
PRIOR FILING DATE: 2001-10-04  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 61  
TYPE: PRT  
ORGANISM: Buthus occitanus  
S-10-264-480-7

Alignment Scores:  
red. No.: 0.932 Length: 61  
core: 71.50 Matches: 24  
Percent Similarity: 55.26% Conservative: 7  
Best Local Similarity: 36.84% Mismatches: 12  
Query Match: 26.78% Indels: 5  
DB: 15 Gaps: 2

US-09-673-274B-2\_COPY\_1\_132 (1-132) x US-10-264-480-7 (1-61):

QY 16 TCCTCGGTGGGGTGGTGTGAACACTTCCGATTGCAACGGTGGTGAAGGAGG 75  
|||||  
Db 13 SerCysLeuPheGly-----AsnGluGlyCysAspLysGluCysLysAlaTyr 28  
|||  
QY 76 GGTACAAAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 129  
|||||  
Db 29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45  
|||

RESULT 12  
US-10-264-480-8  
Sequence 8, Application US/10264480  
Publication No. US20030113892A1  
GENERAL INFORMATION:  
APPLICANT: Hammock, Bruce D.  
APPLICANT: Inceoglu, Bora  
TITLE OF INVENTION: ISOLATED POLYPEPTIDES AND COMPOSITIONS  
TITLE OF INVENTION: FROM THE VENOM OF P. TRANSVAALICUS AND METHODS OF USE  
FILE REFERENCE: UCAL256  
CURRENT APPLICATION NUMBER: US/10/264,480  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: 60/393,070  
PRIOR FILING DATE: 2002-06-28  
PRIOR APPLICATION NUMBER: 60/327,602  
PRIOR FILING DATE: 2001-10-04  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 61  
TYPE: PRT  
ORGANISM: Buthus occitanus  
S-10-264-480-8

Alignment Scores:  
red. No.: 0.932 Length: 61  
core: 71.50 Matches: 14  
Percent Similarity: 55.26% Conservative: 7  
Best Local Similarity: 36.84% Mismatches: 12  
Query Match: 26.78% Indels: 5  
DB: 15 Gaps: 2

US-09-673-274B-2\_COPY\_1\_132 (1-132) x US-10-264-480-8 (1-61)

QY 16 TCCTCGGTGGGGTGGTGTGAACACTTCCGATTGCAACGGTGGTGAAGGAGG 75  
|||||  
Db 13 SerCysLeuPheGly-----AsnGluGlyCysAspLysGluCysLysAlaTyr 28  
|||  
QY 76 GGTACAAAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 129  
|||||  
Db 29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45  
|||

RESULT 13  
US-10-178-213-411  
Sequence 411, Application US/10178213  
Publication No. US20030041348A1  
GENERAL INFORMATION:  
APPLICANT: Simmons, Carl R.  
APPLICANT: Navarrio Acevedo, Pedro A.  
APPLICANT: Harvell, Leslie  
APPLICANT: Cahoon, Rebecca  
APPLICANT: McCutchen, Billy Fred  
APPLICANT: Lu, Albert  
APPLICANT: Herrmann, Rafael  
APPLICANT: Wong, James  
TITLE OF INVENTION: Defensin Polynucleotides and Methods of  
TITLE OF INVENTION: Use  
FILE REFERENCE: 35718/246703  
CURRENT APPLICATION NUMBER: US/10/178,213  
CURRENT FILING DATE: 2002-06-21  
PRIOR APPLICATION NUMBER: 60/300,152  
PRIOR FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: 60/300,241  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 469  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 411  
LENGTH: 47  
TYPE: PRT  
ORGANISM: Tropaeolum majus  
US-10-178-213-411

Alignment Scores:  
red. No.: 1.05 Length: 47  
core: 71.00 Matches: 14  
Percent Similarity: 50.00% Conservative: 5  
Best Local Similarity: 36.84% Mismatches: 13  
Query Match: 26.59% Indels: 6  
DB: 15 Gaps: 2

US-09-673-274B-2\_COPY\_1\_132 (1-132) x US-10-178-213-411 (1-47)

QY 13 GGTTCGTGGTGGGTCCTGTGAACACTTCCGATTGCAACGGTGGTGAAGG 72  
|||||  
Db 12 GlyProCysVal-----SerLysThrAsnCysAlaSerValCysLysTr 26  
|||  
QY 73 AGCGTTACAAAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 126  
|||||  
Db 27 GluGlyPheTyrGlyGlyHscysArgGlyPheArgHis--ArgCysPheCys 43  
|||||

RESULT 14  
US-10-178-213-410  
Sequence 410, Application US/10178213  
Publication No. US20030041348A1  
GENERAL INFORMATION:  
APPLICANT: Simmons, Carl R.  
APPLICANT: Navarrio Acevedo, Pedro A.  
APPLICANT: Harvell, Leslie  
APPLICANT: Cahoon, Rebecca  
APPLICANT: McCutchen, Billy Fred  
APPLICANT: Lu, Albert  
APPLICANT: Herrmann, Rafael  
APPLICANT: Wong, James  
TITLE OF INVENTION: Defensin Polynucleotides and Methods of  
TITLE OF INVENTION: Use  
FILE REFERENCE: 35718/246703  
CURRENT APPLICATION NUMBER: US/10/178,213  
CURRENT FILING DATE: 2002-06-21  
PRIOR APPLICATION NUMBER: 60/300,152  
PRIOR FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: 60/300,241  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 469  
SOFTWARE: FastSeq for Windows Version 4.0

Search completed: October 16, 2003, 17:34:08  
Job time : 28.4086 secs

```

; SEQ ID NO 410
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Tropaeolum majus
US-10-178-213 410

Alignment Scores:
  Pred. No.: 1.1      Length: 77
  Score: 71.00      Matches: 14
  Percent Similarity: 50.00%      Conservative: 5
  Best Local Similarity: 36.84%      Mismatches: 13
  Query Match: 26.59%      Indels: 6
  DB: 15      Gaps: 2

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-178-213-410 (1-77)

CY 13 GGTTCCTGCGTGGGCTGTGAACCTACACTTCGATTCGAACGGTGAATGCAAGAGG 72
Db 42 GlyProCysVal-----SerLysThrAsnCysAlaSerValCysLysThr 56
CY 73 AGGGGTTACAGGGTGGCTGACCTGGCTTCCTTCGCTAACGTCGAACCTGGTGGTC 126
Db 57 GluGlyPheTyGlyGlyHisCysArgGlyPheArgHis---ArgCysPheCys 73

RESULT 15
US-10-178-213 387
; Sequence 387, Application US/10:78213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro Acevedo, Pedro A.
; APPLICANT: Harvell, Leslie
; APPLICANT: Cahoon, Rebecca
; APPLICANT: McCutchen, Billy Fred
; APPLICANT: Lo, Albert
; APPLICANT: Herrmann, Rafael
; APPLICANT: Wong, James
; TITLE OF INVENTION: Defensin Polynucleotides and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 35718/246703
; CURRENT APPLICATION NUMBER: US/10/178,213
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/300,152
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/300,241
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 387
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Glycine max
US-10-178-213 387

Alignment Scores:
  Pred. No.: 1.22      Length: 47
  Score: 70.50      Matches: 10
  Percent Similarity: 59.26%      Conservative: 6
  Best Local Similarity: 37.04%      Mismatches: 10
  Query Match: 26.40%      Indels: 1
  DB: 15      Gaps: 1

US-09-673-274B-2_COPY_1_132 (1-132) x US-10-178-213-387 (1-47)

CY 46 TCCGATTGCAACGGTGAAGAGAGGGGTTAAAGGGTGGTCACTGGGTCCTTC 105
Db 28 ThrAsnCysGlySerValCysArgThrGluGlyPheSerCysGlyHisCysArgGlyPhe 37
CY 106 GCTAACGTGAACCTGCTGGTGC 126
Db 38 SerArg---ArgCysPheCys 43
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

M nucleic - protein search, using frame\_plus\_n2p mode:

un on: October 16, 2003, 17:06:37 ; Search time 13.561 Seconds  
(without alignments)  
849,038 Million cell updates/sec

file: US-09-673-274B-2\_COPY\_1\_132

effect score: 267

sequence: 1 gataagcttctcggtctctg.....tgactgctggtcgagact 132

coring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

searched: 328717 seqs, 42110858 residues

total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODEL=frame+ n2p.model -DEV=xlp  
Q=/cgn2\_1/USPTO\_spool/US09673274/runat\_16102003\_170628\_2580/app\_query.fasta\_1.654  
DB=Issued Patents AA -OPMT-fastan -SUFFIX=ra1 -MINMATCH=0.1 -LOOPCL=0  
LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
MODE=LOCAL -OUTPMT=pro -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
USER=USC9673274 -CGN\_1\_52@runat\_16102003\_170628\_2580 -NCPU=6 -ICPU3  
NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/6C-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80.5	30.1	65	1	US-08-435-040-2
2	80.5	30.1	65	3	US-09-020-216-2
3	76.5	28.7	61	1	US-08-682-485A-25
4	76.5	28.7	61	1	US-08-451-472-4
5	76.5	28.7	61	2	US-08-933-314-25
6	74.5	27.9	44	4	US-09-480-251-4
7	74.5	27.9	69	4	US-09-480-251-6
8	74.5	27.9	70	4	US-09-480-251-2
9	74	27.7	64	4	US-09-403-343B-25
C 10	68	28.9	147	4	US-09-252-991A-16810
C 11	68	28.9	149	4	US-09-252-991A-32270
12	67	25.1	60	1	US-08-451-472-6

13	67	25.1	67	1	US-09-435-040-3	Sequence 3, Appl:
14	67	25.1	67	3	US-09-020-216-3	Sequence 3, Appl:
15	64.5	24.2	80	3	US-08-952-383A-14	Sequence 14, Appl:
16	64.5	24.2	80	4	US-08-970-264A-27	Sequence 27, Appl:
17	63.5	23.8	61	1	US-08-682-485A-26	Sequence 26, Appl:
18	63.5	23.8	61	2	US-08-933-314-26	Sequence 26, Appl:
C 19	63	26.8	631	4	US-09-252-991A-20063	Sequence 20063, A
20	62	23.2	314	2	US-08-460-309-19	Sequence 19, Appl:
21	62	23.2	314	2	US-08-125-077-19	Sequence 19, Appl:
22	61.5	23.0	74	4	US-09-442-631-4	Sequence 4, Appl:
23	61.5	23.0	75	1	US-08-289-458-2	Sequence 2, Appl:
24	61.5	23.0	75	3	US-08-761-549-2	Sequence 2, Appl:
25	61.5	23.0	75	3	US-09-127-646-2	Sequence 2, Appl:
C 26	61.5	26.2	288	4	US-09-252-991A-31433	Sequence 31433, A
C 27	61.5	26.2	301	3	US-09-231-227-4	Sequence 4, Appl:
C 28	61.5	26.2	301	4	US-09-768-585-4	Sequence 4, Appl:
C 29	61	26.2	296	4	US-09-252-991A-32162	Sequence 32162, A
30	60.5	22.7	68	3	US-09-053-021-1	Sequence 1, Appl:
31	60.5	22.7	81	3	US-09-053-021-4	Sequence 4, Appl:
32	60.5	22.7	91	3	US-09-053-021-9	Sequence 9, Appl:
33	59	22.1	359	4	US-09-699-266A-11	Sequence 11, Appl:
34	58.5	21.9	67	3	US-09-053-021-2	Sequence 2, Appl:
35	58	21.7	54	1	US-08-656-318A-1	Sequence 1, Appl:
36	58	21.7	54	2	US-08-956-459-1	Sequence 1, Appl:
37	58	21.7	62	1	US-08-451-472-5	Sequence 5, Appl:
C 38	57.5	24.5	103	4	US-09-252-991A-22787	Sequence 22787, A
C 39	57.5	24.5	518	4	US-09-252-991A-25967	Sequence 25967, A
40	57.5	21.5	2556	1	US-08-185-432-17	Sequence 17, Appl:
41	57.5	21.5	2556	1	US-08-083-590A-20	Sequence 20, Appl:
42	57.5	21.5	2556	3	US-08-532-384-20	Sequence 20, Appl:
43	57.5	21.5	2556	4	US-08-999-232-2	Sequence 2, Appl:
C 44	57	24.3	145	4	US-09-252-991A-22334	Sequence 22334, A
C 45	57	24.3	526	4	US-09-252-991A-23462	Sequence 23462, A

#### ALIGNMENTS

RESULT 1  
US-08-435-040-2  
; Sequence 2, Application: US/09435040  
; Patent No. 5756340  
; GENERAL INFORMATION:  
; APPLICANT: Hammock, Bruce D.  
; APPLICANT: Hermann, Rafael  
; TITLE OF INVENTION: Insect Control With Multiple Toxins  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Majestic, Parsons, Siebert & Hsue  
; STREET: Four Embarcadero Center, Suite 1450  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 94111-4121  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/435,040  
; FILING DATE: 08-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Siebert, J. Suzanne  
; REGISTRATION NUMBER: 28,758  
; REFERENCE/DOCKET NUMBER: 2500.078US0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 362-5556  
; TELEFAX: (415) 362-5418  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 65 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-09-673-274b-2

Alignment Scores:  
Pred. No.: 0.0127 Length: 65  
Score: 80.50 Matches: 17  
Percent Similarity: 57.14% Conservative: 7  
Best Local Similarity: 40.48% Mismatches: 9  
Query Match: 30.15% Indels: 9  
DB: 1 Gaps: 4

US-09-673-274B-2\_COPY\_1\_132 (1-132) x US-08-435-040-2 (1-65)

QY 16 TCCTGGCTGG-----GGTGTGTGACTTCCGATTGCAACGGTGAGTGAAG 69  
Db 12 AsnCysValThrCysGlyAlaAsnSerTyr-----CysAsnThrGluCysThr 28  
QY 70 AGGAGGGTTACAGGGTGGTCACTGC-----GGTCTTCGCTAACTGAACTGC 120  
Db 29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpPheGlyLysTyrGlyAsnAla---Cys 47  
QY 121 TGGTGC 126  
Db 48 TrpCys 49

## RESULT 2

US-09-020-216-2  
Sequence 2, Application US/09020216  
Patent No. 6:62430  
GENERAL INFORMATION:  
APPLICANT: Hammock, Bruce D.  
APPLICANT: Herrmann, Rafael  
APPLICANT: Moskowitz, Haim  
TITLE OF INVENTION: Insect Control With Multiple Toxins  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Majestic, Parsons, Siebert & Hsue  
STREET: Four Embarcadero Center, Suite 1450  
CITY: San Francisco  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 94111-4121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/020,216  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,040  
FILING DATE: 08-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Siebert, J Suzanne  
REGISTRATION NUMBER: 28,758  
REFERENCE/DOCKET NUMBER: 2500,078USU  
TELEPHONE: (415) 362-5556  
TELEFAX: (415) 362-5418  
TELEX: 278638 MGPS  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 65 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-09-020-216-2

Alignment Scores:  
Pred. No.: 0.0127 Length: 65  
Score: 80.50 Matches: 17  
Percent Similarity: 57.14% Conservative: 7  
Best Local Similarity: 40.48% Mismatches: 9  
Query Match: 30.15% Indels: 9  
DB: 1 Gaps: 4

US-09-673-274B-2\_COPY\_1\_132 (1-132) x US-09-020-216-2 (1-65)

QY 16 TCCTGGCTGG-----GGTGTGTGAACTACTTCCGATTGCAACGGTGAGTGAAG 69  
Db 12 AsnCysValThrCysGlyAlaAsnSerTyr-----CysAsnThrGluCysThr 28  
QY 70 AGGAGGGTTACAGGGTGGTCACTGC-----GGTCTTCGCTAACTGAACTGC 120  
Db 29 LysAsnGlyAlaGluSerGlyTyrCysGlnTrpPheGlyLysTyrGlyAsnAla---Cys 47  
QY 121 TGGTGC 126  
Db 48 TrpCys 49

## RESULT 3

US-08-682-485A-25  
Sequence 25, Application US/08682485A  
Patent No. 5763568  
GENERAL INFORMATION:  
APPLICANT: ATKINSON, RONALD K  
APPLICANT: HOWDEN, MERLIN E.H.  
APPLICANT: TYLER, MARGARET I  
APPLICANT: VONARX, EDWARD J  
TITLE OF INVENTION: Insecticidal Toxins Derived From Spiders  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Zeneca, Inc.  
STREET: 1200 South 47th Street  
CITY: Richmond  
STATE: California  
COUNTRY: USA  
ZIP: 94804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/682,485A  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/682,485  
FILING DATE: 17-JULY-1996  
APPLICATION NUMBER: AUS/08/256,933  
FILING DATE: 27-JULY-1994  
APPLICATION NUMBER: WO 93/15108  
FILING DATE: 29-JAN-1993  
APPLICATION NUMBER: AU P0722  
FILING DATE: 31-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Shaw, Melissa A.  
REGISTRATION NUMBER: 38,301  
REFERENCE/DOCKET NUMBER: PPD 5099/D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 510-231-1542  
TELEFAX: 510-231-1112  
INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 61 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Leturus quinquestriatus quinquestriatus

US-08-682-485A-25

Alignment Scores:

Pred. No.: 0.0418 Length: 61  
 Score: 76.50 Matches: 15  
 Percent Similarity: 55.26% Conservative: 6  
 Best Local Similarity: 39.47% Mismatches: 12  
 Query Match: 28.65% Indels: 5  
 DB: 1 Gaps: 2

US-09-673-274B-2\_COPY\_1\_132 (1-132) x US-08-682-485A-25 (1-61)

QY 16 TCCTGGTGGGTGCTGTGAACACTACACTCCGATTGCAACGGTGAGTGCAGAGGAGG 75  
 |||||:|||||  
 Db 13 SerCysLeuPheGly-----AsnGluGlyCysAsnLysGluCysLysSerTyr 28  
 QY 76 GGTACAAAGGGTGCTACCTGCGGTTCTCTTCCCTAACGTGAACGTGCTGTCGCGAG 129  
 |||||:|||||  
 Db 29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 4

US-08-451-472-4

Sequence 4, Application US/08451472

Patent No. 570192

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: BIOLOGICAL CONTROL AGENTS

NUMBER OF SEQUENCES: 73

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon & Vanderhye PC

STREET: 8th Floor, 1100 No. 5770192th Glebe Road

CITY: Arlington

STATE: Virginia

COUNTRY: USA

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/451,472

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/117,125

FILING DATE: 24-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: Mary J Wilson

REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 117-187

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816 4000

TELEFAX: (703) 816 4100

TELEX: 200797 NIXN UR

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 61 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-451-472-4

Alignment Scores:

Pred. No.: 0.0418 Length: 61  
 Score: 76.50 Matches: 15  
 Percent Similarity: 55.26% Conservative: 6  
 Best Local Similarity: 39.47% Mismatches: 12  
 Query Match: 28.65% Indels: 5  
 DB: 2 Gaps: 2

US-09-673-274B-2\_COPY\_1\_132 (1-132) x US-08-451-472-4 (1-61)

QY 16 TCCTGGTGGGTGCTGTGAACACTACACTCCGATTGCAACGGTGAGTGCAGAGGAGG 75  
 |||||:|||||  
 Db 13 SerCysLeuPheGly-----AsnGluGlyCysAsnLysGluCysLysSerTyr 28  
 QY 76 GGTACAAAGGGTGCTACCTGCGGTTCTCTTCCCTAACGTGAACGTGCTGTCGCGAG 129  
 |||||:|||||  
 Db 29 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 45

RESULT 5

US-08-933-314-25

Sequence 25, Application US/08933314

Patent No. 5959192

GENERAL INFORMATION:

APPLICANT: ATKINSON, RONALD K

ADDRESSEE: HOWDEN, MERLIN F.H.

STREET: TYLER, MARGARET

CITY: VONARX, EDWARD J

TITLE OF INVENTION: Insecticidal Toxins Derived From

TITLE OF INVENTION: Funnel Web (Atrax or Hadronyche Spiders:

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Zeneca, Inc.

STREET: 1200 South 47th Street

CITY: Richmond

STATE: California

COUNTRY: USA

ZIP: 94804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/933,314

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/682,495

FILING DATE: 17-JULY-1996

APPLICATION NUMBER: US/08/256,933

FILING DATE: 27-JULY-1994

APPLICATION NUMBER: WO 93/15108

FILING DATE: 29-JAN-1993

APPLICATION NUMBER: AU PLC722

FILING DATE: 31-JAN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Shaw, Melissa A.

REGISTRATION NUMBER: 38,301

REFERENCE/DOCKET NUMBER: PPD 5099/D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 510-231-1542

TELEFAX: 510-231-1112

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 61 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Leturus quinquestriatus quinquestriatus

US-08-933-314-25

```

Alignment Scores:
Pred. No.: 0.0418 Length: 61
Score: 76.50 Matches: 15
Percent Similarity: 55.26% Conservative: 6
Best Local Similarity: 39.47% Mismatches: 12
Query Match: 28.65% Indels: 5
DB: 2 Gaps: 2

US-09-673-274B-2_COPY_1_132 (1-132) x US-08-933-314-25 (1-61)
QY 16 TCTTCGGTGGGTGGTGTGAACACACTCCGATTCGACAGCGGTGAGTGCACAGGAGG 75
Dc 13 SerCysLeuPheGly-----AsnGluG:YcysAsnLysGluCysCysSerTyr 28
QY 76 GGTACAGGGGTGGTGTGCTGCTTCCTTCCTGCTAACGTGAACCTGCTGGTGGCGAG 129
Dc 29 GlyCysSerTyrG:YTyrcysTrpThrTrp GlyLeuAlaCysTrpCysGlu 45

RESULT 6
US-09-480-251-4
; Sequence 4, Application US/09480251
; Patent No. 6465719
; GENERAL INFORMATION:
; APPLICANT: DeRose, Richard
; APPLICANT: Freyssinet, Georges
; APPLICANT: Hoffman, Jules
; TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
; TITLE OF INVENTION: Vector Containing It And Production Of Disease-Resistant
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: A32889-PCT-USA-A
; CURRENT FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/FR98/01462
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: FRANCE 97/09,115
; PRIOR FILING DATE: 1997-07-11
; PRIOR APPLICATION NUMBER: FRANCE 9709,663
; PRIOR FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-480-251-4
Alignment Scores:
Pred. No.: 0.073 Length: 44
Score: 74.50 Matches: 15
Percent Similarity: 42.50% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 16
Query Match: 27.90% Indels: 7
DB: 4 Gaps: 3

US-09-673-274B-2_COPY_1_132 (1-132) x US-09-480-251-4 (1-44)
QY 13 GGTTCCTGC---GTGTGGGTGCTGTGAACACACTTCGATTCGACACGGTGAAGTGCAG 69
Dc 9 GlyProCysAlaValTrpAsp-----AsnGluThrCysArgValCysCys 24
QY 70 AGGAGGGGTACAGGGTGGTGTGCTGCTTCCTTCCTGCTAACGTGAACCTGCTGGTGGCGAG 129
Dc 25 GluGluGlyArgSerGlyHisCysSer-----ProSerLeuLysCysTrpCysGlu 42

RESULT 7
US-09-480-251-6
; Sequence 6, Application US/09480251
; Patent No. 6465719
; GENERAL INFORMATION:
; APPLICANT: DeRose, Richard
; APPLICANT: Freyssinet, Georges
; APPLICANT: Hoffman, Jules
; TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
; TITLE OF INVENTION: Vector Containing It And Production Of Disease-Resistant
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: A32889-PCT-USA-A
; CURRENT FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/FR98/01462
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: FRANCE 97/09,115
; PRIOR FILING DATE: 1997-07-11
; PRIOR APPLICATION NUMBER: FRANCE 9709,663
; PRIOR FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-480-251-6
Alignment Scores:
Pred. No.: 0.0775 Length: 69
Score: 74.50 Matches: 15
Percent Similarity: 42.50% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 16
Query Match: 27.90% Indels: 7
DB: 4 Gaps: 3

US-09-673-274B-2_COPY_1_132 (1-132) x US-09-480-251-6 (1-69)
QY 13 GGTTCCTGC---GTGTGGGTGCTGTGAACACACTTCGATTCGACACGGTGAAGTGCAG 69
Dc 34 GlyProCysAlaValTrpAsp-----AsnGluThrCysArgValCysLys 49
QY 70 AGGAGGGGTACAGGGTGGTGTGCTGCTTCCTTCCTGCTAACGTGAACCTGCTGGTGGCGAG 129
Dc 50 GluGluGlyArgSerGlyHisCysSer-----ProSerLeuLysCysTrpCysGlu 67

RESULT 8
US-09-480-251-2
; Sequence 2, Application US/09480251
; Patent No. 6465719
; GENERAL INFORMATION:
; APPLICANT: DeRose, Richard
; APPLICANT: Freyssinet, Georges
; APPLICANT: Hoffman, Jules
; TITLE OF INVENTION: Chimeric Gene Encoding Drosomycin,
; TITLE OF INVENTION: Vector Containing It And Production Of Disease-Resistant
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: A32889-PCT-USA-A
; CURRENT FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/FR98/01462
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: FRANCE 97/09,115
; PRIOR FILING DATE: 1997-07-11
; PRIOR APPLICATION NUMBER: FRANCE 9709,663
; PRIOR FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-480-251-2
Alignment Scores:
Pred. No.: 0.0777 Length: 70
Score: 74.50 Matches: 15
Percent Similarity: 42.50% Conservative: 2
Best Local Similarity: 37.50% Mismatches: 16
Query Match: 27.90% Indels: 7

```



```
DB: 4 Gaps: 3
JS-09-673-274B-2_COPY_1_132 (1-132) x US-09-480-251-2 (1-70)
Y 13 GGTCTCTCTC---GTCTGGGGTGTCTGTGAATACACATTCGCGATTGCAACGGTGAAGTGAAG 69
b 35 GlyProCysAlaValTipAsp-----AsnGluThrCysArgValCysLys 50
Y 70 AGGAGGGGTACAGGGTGTCTACATCGCGTTCCTTCGTAACGTGAATGCTGTGTGGAG 129
b 51 GluGluGlyArgSerSerGlyHisCysSer-----ProSerLeuLysCysTrpCysGlu 68
RESULT 9
JS-09-403-343B-25
Sequence 25, Application US/09403343B
Patent No. 655091
GENERAL INFORMATION:
APPLICANT: JOLIVET-REYNAUD, COLETTE
APPLICANT: PERRON, HERVE
APPLICANT: MANDRAND, BERNARD
TITLE OF INVENTION: POLYPEPTIDE CAPABLE OF REACTING WITH ANTIBODIES OF
TITLE OF INVENTION: PATIENTS SUFFERING FROM MULTIPLE SCLEROSIS AND USES
FILE REFERENCE: 104574
CURRENT APPLICATION NUMBER: US/09/403,343B
CURRENT FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: FR/97/05679
PRIOR FILING DATE: 1997-04-29
PRIOR APPLICATION NUMBER: FR/97/16870
PRIOR FILING DATE: 1997-12-31
PRIOR APPLICATION NUMBER: PCT/FR98/00870
PRIOR FILING DATE: 1998-04-29
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 25
LENGTH: 64
TYPE: PRT
ORGANISM: MSRV
JS-09-403-343B-25
Alignment Scores:
Pred. No.: 0.0892 Length: 64
Score: 74.00 Matches: 12
Percent Similarity: 59.26% Conservative: 4
Best Local Similarity: 44.44% Mismatches: 9
Query Match: 27.72% Indels: 2
DB: 4 Gaps: 1
JS-09-673-274B-2_COPY_1_132 (1-132) x US-09-403-343B-25 (1-64)
Y 52 TGCAACGGTGAAGGAGGGGTACAGGGTGTCTACAGGGTGTCTCTCTC----- 105
b 20 CysAspGlyLeuCysLysAsnGlyAlaLysSerGlySerCysGlyPheLeuValPro 39
Y 106 GCTAACGTGAATGCTGTGTC 126
b 40 SerGlyLeuAlaCysTrpCys 46
RESULT 10
US-09-252-991A-16810
Sequence 16810, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
```

```
; SEQ ID NO 16810
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16810
Alignment Scores:
Pred. No.: 0.605 Length: 147
Score: 68.00 Matches: 15
Percent Similarity: 59.38% Conservative: 4
Best Local Similarity: 46.84% Mismatches: 11
Query Match: 28.94% Indels: 2
DB: 4 Gaps: 2
US-09-673-274B-2_COPY_1_132 (1-132) x US-09-252-991A-16810 (1-147)
Y 106 CGAAGGACCGGAGTGAC---CACCTTTGTAAACCTCTCTTGGCACTACCGTTGCAAT 50
b 45 ArgArgAsnProAsnProPheHisProCys---ProArgAsnCysAlaGlyArgSerAsn 63
Y 49 CGGAAGTGTAGTTCACAGCAGCCGACACACGACGAAC 14
b 64 GlnProCysLeuValProAsnProProArgGAsn 75
RESULT 11
US-09-252-991A-32270
Sequence 32270, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32270
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32270
Alignment Scores:
Pred. No.: 0.606 Length: 149
Score: 68.00 Matches: 17
Percent Similarity: 38.89% Conservative: 4
Best Local Similarity: 31.48% Mismatches: 13
Query Match: 28.94% Indels: 20
DB: 4 Gaps: 2
US-09-673-274B-2_COPY_1_132 (1-132) x US-09-252-991A-32270 (1-149)
Y 112 CGTAGGAGGAGCGGAGTGACACCGCT-----TGTAACCC 24
b 7 ArgTrpArgArgAsnSerSerProSerProAlaSerArgArgArgArgCysSerPhe 26
Y 73 TCC-----TCTTGCACTCACCGTTGC 53
b 27 SerArgSerProThrGlyArgTyrThrArgAlaProTtpArgCysCysArgArgArgCys 46
Y 52 AATCGGAAGTGTATTCACAGCAGCCACCCACACGACGACGACCGA 11
b 47 ArgArgArgCysArgProArgProThrSerLysAlaArg 60
RESULT 12
US-08-451-472-6
Sequence 6, Application US/08451472
Patent No. 5770192
GENERAL INFORMATION:
```

1 APPLICANT:  
2 TITLE OF INVENTION: BIOLOGICAL CONTROL AGENTS  
3 NUMBER OF SEQUENCES: 73  
4 CORRESPONDENCE ADDRESS:  
5 ADDRESSEE: Nixon & Vanderhye PC  
6 STREET: 8th Floor, 1100 No. 5770192th Glebe Road  
7 CITY: Arlington  
8 STATE: Virginia  
9 COUNTRY: USA  
10 ZIP: 22201-4714  
11  
12 COMPUTER READABLE FORM:  
13 MEDIUM TYPE: Floppy disk  
14 COMPUTER: IBM PC compatible  
15 OPERATING SYSTEM: PC-DOS/MS-DOS  
16 CORRESPONDENCE ADDRESS:  
17 ADDRESSEE: Nixon & Vanderhye PC  
18 STREET: 8th Floor, 1100 No. 5770192th Glebe Road  
19 CITY: Arlington  
20 STATE: Virginia  
21 COUNTRY: USA  
22 ZIP: 22201-4714  
23  
24 MEDIUM TYPE: Floppy disk  
25 COMPUTER: IBM PC compatible  
26 OPERATING SYSTEM: PC-DOS/MS-DOS  
27 SOFTWARE: Patent In Release #1.0, Version #1.25  
28 CURRENT APPLICATION DATA:  
29 APPLICATION NUMBER: US/08/451,472  
30 FILING DATE:  
31 CLASSIFICATION: 435  
32 PRIOR APPLICATION NUMBER: US 08/117,125  
33 FILING DATE: 24-NOV-1993  
34 ATTORNEY/AGENT INFORMATION:  
35 NAME: Mary J. Wilson  
36 REGISTRATION NUMBER: 32,955  
37 REFERENCE/DOCKET NUMBER: 117-187  
38 TELEPHONE: (703) 816 4003  
39 TELEFAX: (703) 816 4100  
40 TELEX: 200797 NIXN UR  
41 INFORMATION FOR SEQ ID NO: 6:  
42 SEQUENCE CHARACTERISTICS:  
43 LENGTH: 60 amino acids  
44 TYPE: amino acid  
45 TOPOLOGY: linear  
46 MOLECULE TYPE: peptide  
47 US-08 451-472-6

Alignment Scores:  
Pred. No.: 0.726 Length: 60  
Score: 67.00 Matches: 13  
Percent Similarity: 56.76% Conservative: 8  
Best Local Similarity: 35.14% Mismatches: 10  
Query Match: 25.09% Indels: 6  
Gaps: 2

US-09 673-2745-2\_COPY\_1\_132 (1-132) x US-08 451-472-6 (1-60)

QY 19 TCGTGTGGGGTGGTGTGTAACACTTCCGATTGCAACGGTGAGTCAAGAGGGGGT 78  
Db 14 CysLeuIleGly-----AsnGluCysAspLysGluCysLysAlaTyrGly 28  
QY 79 TACAGGGTGGTCACTGGCGGTTCCTCGGTACAGTGAATGTTGGTGGAG 129  
Db 29 GlySerTyrGlyTyrCysThrTrp---GlyLeuAlaCysTrpCysGlu 44

RESULT 13  
US-08-435-040-3  
Sequence 3, Application US/08435040  
Patent No. 5756340  
GENERAL INFORMATION:  
APPLICANT: Hammock, Bruce D.  
APPLICANT: Herrmann, Rafael  
APPLICANT: Moskowitz, Haim  
TITLE OF INVENTION: Insect Control With Multiple Toxins  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Majestic, Parsons, Siebert & Hsue  
STREET: Four Embarcadero Center, Suite 1450  
CITY: San Francisco  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 94111-4121

1 COMPUTER READABLE FORM:  
2 MEDIUM TYPE: Floppy disk  
3 COMPUTER: IBM PC compatible  
4 OPERATING SYSTEM: PC-DOS/MS-DOS  
5 SOFTWARE: Patent In Release #1.0, Version #1.25  
6 CURRENT APPLICATION DATA:  
7 APPLICATION NUMBER: US/08/435,040  
8 FILING DATE: 08-MAY-1995  
9 CLASSIFICATION: 514  
10 ATTORNEY/AGENT INFORMATION:  
11 NAME: Siebert, J. Suzanne  
12 REGISTRATION NUMBER: 28,758  
13 REFERENCE/DOCKET NUMBER: 2500.078JUSC  
14 TELEPHONE: (415) 362-5556  
15 TELEFAX: (415) 362-5428  
16 TELEX: 278638 MOPS  
17 INFORMATION FOR SEQ ID NO: 3:  
18 SEQUENCE CHARACTERISTICS:  
19 LENGTH: 67 amino acids  
20 TYPE: amino acid  
21 TOPOLOGY: linear  
22 MOLECULE TYPE: peptide  
23 HYPOTHETICAL: NO  
24 ANTI-SENSE: NO  
25 FRAGMENT TYPE: N-terminal  
26 US-08-435-040-3

Alignment Scores:  
Pred. No.: 0.736 Length: 67  
Score: 67.00 Matches: 13  
Percent Similarity: 46.15% Conservative: 5  
Best Local Similarity: 33.33% Mismatches: 19  
Query Match: 25.09% Indels: 2  
Gaps: 1

US-09-673-274B-2\_COPY\_1\_132 (1-132) x US-08-435-040-3 (1-67)

QY 16 TCCTGGTGGGGTGGTGTGTAACACTTCCGATTGCAACGGTGAGTCAAGAGGAGG 75  
Db 12 AsnCysValTyrHisCysPheProGlySerProGlyCysAspThrLeuCysLysGlyAsp 31  
QY 76 GGTACAGGGTGGTCACTGGCGGT-----TCCTTCGCTAACGTGAACACTGCTGGTGC 126  
Db 32 GlyAlaSerSerGlyHisCysGlyPheLysGluGlyHisGlyLeuAlaCysTrpCys 50

RESULT 14  
US-09-020-216-3  
Sequence 3, Application US/09020216  
Patent No. 6162430  
GENERAL INFORMATION:  
APPLICANT: Hammock, Bruce D.  
APPLICANT: Herrmann, Rafael  
APPLICANT: Moskowitz, Haim  
TITLE OF INVENTION: Insect Control With Multiple Toxins  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Majestic, Parsons, Siebert & Hsue  
STREET: Four Embarcadero Center, Suite 1450  
CITY: San Francisco  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 94111-4121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/020,216  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/435,040  
FILING DATE: 08-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Siebert, J. Suzanne  
REGISTRATION NUMBER: 28,758  
REFERENCE/DOCKET NUMBER: 2500.078USO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 362-5556  
TELEFAX: (415) 362-5418  
TELEX: 278638 MGPS  
INFORMATION FOR SEQ ID NO.: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 67 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
JS-09-020-216-3

Alignment Scores:  
Pred. No.: 0.736 Length: 67  
Score: 67.00 Matches: 13  
Percent Similarity: 46.15% Conserv: 5  
Least Local Similarity: 33.33% Mismatches: 19  
Query Match: 25.09% Indels: 2  
DB: 3 Gaps: 1

JS-09-673-274b-2\_COPY\_1\_132 (1-132) x US-09-020-216-3 (1-67)

16 TCCTCGGTGGGGTGTGTAACCTACACTTCGATTCGACGGTGAAGAGGAGG 75  
12 AsnCysValThrHisCysPheProGlyCysAspThrLeuCysGlyAsp 31  
76 GGTACAAAGGGTGTCTACTGCGGT-----TCCTTCGCTAACGTGAACCTGCTGGTGC 126  
32 GlyAlaSerSerGlyHisCysGlyPheGlyGluGlyHisGlyLeuAlaCysTrpCys 50

## RESULT 15

JS-08-952-383A-14  
Sequence 14, Application US/08952383A  
Patent No. 6096304  
GENERAL INFORMATION:

APPLICANT: RECOMBINANT BACULOVIRUS INSECTICIDES

TITLE OF INVENTION: RECOMBINANT BACULOVIRUS INSECTICIDES

NUMBER OF SEQUENCES: 14

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE 3.5 INCH

COMPUTER: IBM

OPERATING SYSTEM: MICROSOFT WINDOWS 3.1

SOFTWARE: MICROSOFT WORD FOR WINDOWS 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/952,383A

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: FLOYD, LINDA AXAMETHY

REGISTRATION NUMBER: 33,692

REFERENCE/DOCKET NUMBER: BA-9063-A

INFORMATION FOR SEQ ID NO.: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 80 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-952-383A-14

## Alignment Scores:

Pred. No.: 1.6 Length: 80

Score: 64.50 Matches: 13

Percent Similarity: 52.63% Conserv: 7

Best Local Similarity: 34.2% Mismatches: 13  
Query Match: 24.16% Indels: 5  
DB: 3 Gaps: 2

US-09-673-274b-2\_COPY\_1\_132 (1-132) x US-08-952-383A-14 (1-80)

16 TCCTCGGTGGGGTGTGTAACCTACACTTCGATTCGACGGTGAAGAGGAGG 75  
32 AlaCysLeuLeuGly-----AsnGluGlyCysAspLysGluCysLysAlaTyr 47  
76 GGTACAAAGGGTGTCTACTGCGGTTCCTTCGCTAACGTGAACCTGCTGGTGCAG 129  
48 GlyGlySerTyrGlyTyrCysTrpThrTrp---GlyLeuAlaCysTrpCysGlu 64

Search completed: October 16, 2003, 17:13:06

Job time : 14.1561 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: October 16, 2003, 17:06:37 : Search time 20.7741 Seconds  
(without alignments)  
1564.690 Million cell updates/sec

Title: US-09-673-274B-2  
Perfect score: 324  
Sequence: 1 gaagacttatcggttcctg.....gaacgtgtagcgatccgg 169

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_n2p.model -DEV=slp  
-Q=/cgn2.1/USPTO.spool/US09673274/runat.16102003.170628.2555/app.query.fasta.1.654  
-DB=PIR\_76 -QFMT=fastan -SUFFIX=prp -MINMATCH=0.1 -LOOPEL=0 -JOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosu62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=ext -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=plc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09673274@cgn.1.1.77@cgnat.16102003.170628.2555 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR\_76:  
1: pir1:  
2: pir2:  
3: pir3:  
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARYES

Result No.	Score	Query Match	Length	ID	Description
1	86	26.5	66	2	A29386 neurotoxin III - s
2	80.5	24.8	65	2	A59222 neurotoxin X - sco
3	79.5	24.5	66	2	A55869 crustacean-specific
4	79	24.4	85	2	S68906 alpha-toxin Bot X:
5	78.5	24.2	84	1	NTSR3A neurotoxin I:1 pre
6	77.5	23.9	65	1	NTSR4L neurotoxin IV - Eg
7	76.5	23.6	61	2	B34123 depressant insect
8	76	23.5	66	1	NTSR4E neurotoxin M14 - 1
9	75.5	23.3	65	1	NTSR1C neurotoxin 1 - bar
10	75	23.1	64	2	JC1321 neurotoxin IV - Sa
11	74.5	23.0	70	2	A55824 drosomycin precurs
12	74	22.8	59	2	D23727 neurotoxin V-5 - b
13	73	22.5	64	1	NTSR5L neurotoxin V - Egy
14	71	21.9	1111	2	T26972 hypothetical prote

15	70.5	21.8	65	1	NTSR1C neurotoxin 3 - bar
16	70.5	21.8	87	2	CNC670 Na+-channel-blocki
17	69.5	21.5	84	1	NTSR2N Na+-channel-blocki
18	68.5	21.1	62	2	C59352 depressant insect
19	68.5	21.1	66	1	NTSR9E neurotoxin X9 - le
20	68.5	21.1	83	1	NTSR1A neurotoxin I precu
21	68	21.0	64	1	NTSR5M neurotoxin V - sco
22	67.5	20.8	67	2	R59352 proteinase inhibit
23	67.5	20.8	67	2	R59352 depressant insect
24	67.5	20.8	67	2	C59352 depressant insect
25	67.5	20.8	83	1	NTSR1E neurotoxin I, prec
26	67.5	20.8	87	2	CNC669 Na+-channel-blocki
27	67	20.7	86	2	CNC671 Na+-channel-blocki
28	66.5	20.5	65	1	NTSR1B neurotoxin I - sco
29	66.5	20.5	64	1	NTSR2C neurotoxin 2 - bar
30	66.5	21.9	873	2	B75514 penicillin-binding
31	66.5	21.9	1791	2	T02345 hypothetical prote
32	66	20.4	3672	2	T23433 hypothetical prote
33	66	20.4	3704	2	T37316 probable laminin a
34	65.5	20.2	67	2	A59306 excitatory insect
35	64.5	19.9	66	2	B23727 neurotoxin V-4 - b
36	64.5	19.9	85	2	A61616 depressant insect
37	64.5	19.9	379	2	H96636 protein FN2.16 f
38	64	19.8	77	2	S30578 protease inhibitor
39	64	19.8	78	2	S52634 gamma-thionin - Pe
40	64	19.8	92	2	S43228 defensin - fruit f
41	64	21.1	1017	2	T18488 hypothetical prote
42	63.5	19.6	85	2	A40472 depressant insect
43	63	19.4	64	2	A35940 neurotoxin alpha f
44	63	19.4	85	2	A39306 alpha insect toxin
45	63	20.8	965	2	S62935 hypothetical prote

## ALIGNMENTS

RESULT 1  
A29386  
neurotoxin III - scorpion (Buthus occitanus)  
C:Species: Buthus occitanus mardochei  
C>Date: 21-May-1988 #sequence, revision 21-May-1988 #text\_change 07-Feb-1997  
C:Accession: A29386  
R:Vargas, O.; Martin, M.F.; Rochat, H.  
Eur. J. Biochem. 162, 589-599, 1987  
A>Title: Characterization of six toxins from the venom of the Moroccan scorpion Buthus  
A:Reference number: A29386; MUD:87161829; PMID:3104036  
A:Accession: A29386  
A:Molecule type: protein  
A:Residues: 1-66 <VAR>  
C:Superfamily: scorpion neurotoxin  
C:Keywords: neurotoxin

Alignment Scores:					
Pred. NO.:	0.0475	Length:	66		
Score:	86.00	Matches:	15		
Percent Similarity:	55.00%	Conservative:	7		
Best Local Similarity:	37.50%	Mismatches:	16		
Query Match:	26.54%	Indels:	2		
DB:	2	Gaps:	1		
US-09-673-274B-2 (1-169) x A29386 (1-66)					
QY	16	TCCTGCGGTGGGGTGGTGTGTAACACTACACTTCGATTGCAACGGTGAAGTCAAGAGGAGG	75		
DB	11	AsnCysValIyrHisCysPheProGlySerSerGlyCysAspThrLeuGlyCysLysGluGlyS	30		
QY	76	GGTTACAAGGTTGGTCTACTCGGTTCTCTC-----GCTAACGTGAACCTGCTGCTGGAG	129		
DB	31	GlyAlaThrSerGlyHisCysGlyPheLeuProGlySerGlyValA.aCysTrpCysAsp	50		
RESULT 2					
A59222					
neurotoxin X - scorpion (Leiurus quinquestriatus)					
C:Species: Leiurus quinquestriatus hebraeus					

```
C>Date: 31-Dec-2001 #sequence_revision 31-Dec-2001 #text_change 19-Apr-2002
C:Accession: A59222
R:Corzo, G.; Escoubas, P.; Nakajima, T.
C:Species: Centruroides limpidus limpidus
A:Submitted to the Protein Sequence Database, April 2000
A:Description: A procedure for the rapid purification of non commercial and new scorpion
A:Reference number: A59222
A:Accession: A59222
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-65 <COR>
C:Superfamily: scorpion neurotoxin
C:Keywords: amidated carboxyl end; neurotoxin; venom
F:13-64,17-37,23 47,21-49/Disulfide bonds: #status predicted
F:65/Modified site: amidated carboxyl end (Arg); #status predicted

Alignment Scores:
Pred. No.: 0.217 Length: 65
Score: 80.50 Matches: 17
Percent Similarity: 57.14% Conservative: 7
Best Local Similarity: 40.48% Mismatches: 9
Query Match: 24.85% Indels: 9
DB: 2 Gaps: 4

US-09-673-274B-2 (1-169) x A59222 (1-65)
QY 16 TCCTCGGTGG-----GGTCTGTGAACACTACCTCCGATTGCAACGGTGAGTGCAG 69
EE 12 AsnCysValTyrThrCysGlyAlaAsnSerTyr-----CysAsnThrGluCysThr 28
QY 20 ACCAGGGGTACAGGGTGGTCACTGC-----GGTCTGTGCTAACTGCAACTGC 120
DB 29 LysAsnGlyAlaGluSerGlyTyrCysGlnTyrPheGlyLysTyrGlyAsnAla---Cys 47
QY 12 TGCTGC 126
DB 48 TrpCys 49

RESULT 3
A59222
Crustacean-specific toxin I - scorpion (Centruroides limpidus)
C:Species: Centruroides limpidus limpidus
C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 17-Mar-1999
C:Accession: A55869
R:Lebreton, P.; Delapierre, M.; Ramirez, A.N.; Salderas, C.; Possani, L.D.
A:Title: Primary and NMR three-dimensional structure determination of a novel crustacean
A:Reference number: A55869; MUID:95244424; PMID:7727365
A:Accession: A55869
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-66 <LEB>
C:Superfamily: scorpion neurotoxin
C:Keywords: disulfide bond; toxin
F:1-2 65,16 41,35-46,29-48/Disulfide bonds: #status experimental

Alignment Scores:
Pred. No.: 0.285 Length: 66
Score: 79.50 Matches: 17
Percent Similarity: 53.85% Conservative: 4
Best Local Similarity: 43.59% Mismatches: 11
Query Match: 24.54% Indels: 7
DB: 2 Gaps: 3

US-09-673-274B-2 (1-169) x A55869 (1-66)
QY 19 TCGTGTGGGTCTGTGTAACACTACCTCCGATTGCAACGGTGAGTGCAGAGGAGGGGT 78
DB 16 CysPheTrpLeuGlyLysAsn-----GluAsnCysAspLysGluCysLysAlaLysAsn 33
QY 79 TACAAGGGT-----GGTCACTCGGTTCTCTCGTTAAAGTGAACCTGCTGGTGCAG 129
DB 34 GlnGlyGlySerTyrGlyTyrCysTyrSerPheAla-----CysTrpCysGlu 49
```

```
RESULT 4
S68906
alpha-toxin Bot XIV precursor - scorpion (Buthus occitanus)
C:Species: Buthus occitanus tunetanus
C:Date: 23-Jul-1997 #sequence_revision 01-Aug-1997 #text_change 16-Jul-1999
C:Accession: S68906
R:Bouhaouala-Zahar, B.; Ducancel, F.; Zencak, I.; ben Khalifa, R.; Borchani, L.; Pei
Edu. J. Biochem. 238, 653-660, 1996
A:Title: A recombinant insect-specific alpha-toxin of Buthus occitanus tunetanus scori
A:Reference number: S68906; MUID:96302228; PMID:8706664
A:Accession: S68906
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-85 <BOU>
A:Cross-references: GB:X92176; NID:gl041277; PIDN:CAA63120.1; PID:gl041278
C:Superfamily: scorpion neurotoxin
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-85/Product: alpha-toxin #status predicted <MAT>

Alignment Scores:
Pred. No.: 0.318 Length: 85
Score: 79.00 Matches: 14
Percent Similarity: 50.00% Conservative: 6
Best Local Similarity: 35.00% Mismatches: 18
Query Match: 24.38% Indels: 2
DB: 2 Gaps: 1

US-09-673-274B-2 (1-169) x S68906 (1-85)
QY 16 TCCTGTGGTGGGTGCTGTGAACACTACCTCCGATTGCAACGGTGAGTGCAGAGGAGG 75
DB 29 AsnCysAlaTyrHisCysLeuLysSerGlyCysAspThrLeuCysLysGluAsn 48
QY 76 GGTACAGGGGTGCTACTCCGCT-----TCCTTCGCTAACGTAACCTGCTGTCGAG 129
DB 49 GlyAlaThrSerGlyHisCysGlyHisLysSerGlyHisGlySerAlaCysTrpCysLys 68

RESULT 5
N5R3A
neurotoxin III precursor - Sahara scorpion
N:Alternate names: mammalian neurotoxin II
C:Species: Androctonus australis (Sahara scorpion)
C:Date: 30-Sep-1979 #sequence_revision 04-Oct-1996 #text_change 18-Jun-1999
C:Accession: C34444; A01740
R:Bougs, P.E.; Rochat, H.; Smith, L.A.
J. Biol. Chem. 264, 19259-19265, 1989
A:Title: Precursors of Androctonus australis scorpion neurotoxins. Structures of prec
A:Reference number: A34444; MUID:90037062; PMID:2808423
A:Accession: C34444
A:Molecule type: mRNA
A:Residues: 1-84 <BOU>
A:Cross-references: GB:M27703; GB:J05102; NID:gl61142; PIDN:AAA29948.1; PID:gl61143
A:Experimental source: clone pCD-634
R:Kopeyar, C.; Martinez, G.; Rochat, H.
Eur. J. Biochem. 94, 639-645, 1979
A:Title: Amino acid sequence of neurotoxin III of the scorpion Androctonus australis I
A:Reference number: A01740; MUID:79148414; PMID:428402
A:Accession: A01740
A:Molecule type: protein
A:Residues: 20-26, 'N', 28-83 <KOP>
C:Superfamily: scorpion neurotoxin
C:Keywords: neurotoxin; venom
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-83/Product: neurotoxin III #status experimental <MAT>
F:31-81,35-53,39-63,43-65/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.: 0.366 Length: 84
Score: 78.50 Matches: 15
Percent Similarity: 53.85% Conservative: 6
Best Local Similarity: 38.46% Mismatches: 13
Query Match: 24.23% Indels: 5
DB: 1 Gaps: 2
```





Db 333 AsnCysThrCysProGlyLeuGupHeSerAspSerAsnAlaSerCysAsp 349

RESULT 15  
NTSR3C  
neurotoxin 3 - bark scorpion  
C:Species: Centruroides sculpuratus (bark scorpion)  
C:Date: 15-Oct-1982 #sequence revision 15-Oct-1982 #text\_change 23-Aug-1986  
C:Accession: A90058; A94470; AC1754  
R:Babin, D.R.; Watt, D.D.; Goos, S.M.; Mlejnek, R.V.  
Arch. Biochem. Biophys. 164, 594-706, 1974  
A:Title: Amino acid sequences of neurotoxic protein variants from the venom of Centruroides sculpuratus (Bask scorpion)  
A:Reference number: A90058; PMID:75163395; PMID:4460885  
A:Accession: A90058

A:Molecule type: protein  
 R:Residues: 1-24, 'NTC', 128-63, 'CS', 'SAB'  
 R:Show, A:; More, J.  
 unpublished results, cited by Pontecilla-Camps, J.C., et al., *Toxicol* 20, 17, 1982  
 A:Reference number: A94470  
 A:Accession: A94470  
 A:Molecule type: protein  
 R:Residues: 1-65 '<RNC>  
 R: Pontecilla-Camps, J.C.; Almasy, R.J.; Suddath, F.L.; Bugg, C.E.  
*Toxicol* 20, 1-7, 1982  
 A:Title: The three-dimensional structure of scorpion neurotoxins.  
 A:Reference number: A94314; MID:92200153; PMID:7880025  
 A:Contents: annotation: X-ray crystallography, 1.8 angstroms; disulfide bonds  
 A:Note: X-ray crystallographic studies were based on the revised sequence shown  
 C:Superfamily: scorpion neurotoxin  
 C:Keywords: neurotoxin; venom  
 F:12-65,16-41,25-46,29-48/Disulfide bonds: #status experimental

Alignment Scores:		
Pred. No.:	3.42	Length: 65
Score:	70.50	Matches: 15
Percent Similarity:	51.22%	Conservative: 6
Best Local Similarity:	36.59%	Mismatches: 13
Query Match:	21.76%	Indels: 7
DB:	1	Gaps: 3
US-09-673-274B-2 (1-169) x NTR3C (1-65)		
Qy	19	TGGGTGGGGTGCTGTGAACACTAC-----ACTTCGATTGCAACGGTGAGTGCAGAGG 72
Db	12	CysLysrtyrGlyCysLeuLysLeuGlyGluAsnGlyCysAspThrGlyCysLysAla 31
Qy	73	AGGGGTTACAAGGTT-----GGTCACGGGTTCCTTCGGTCACTGCAACTGCTGGTGC 126
Db	32	LysAsnGluGlyGlyserTyrGlyTyrCysTyrAlaIleAla-----CysTrpCys 48
Qy	127	GAG 129
Db	49	Glu 49

```
Search completed: October 16, 2003, 17:11:58
Job time : 23.7741 secs
```